

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/52, 9/00, A61K 31/70, C12Q 1/68, C07H 21/00		A1	(11) International Publication Number: WO 95/24489 (43) International Publication Date: 14 September 1995 (14.09.95)
(21) International Application Number: PCT/US95/02816 (22) International Filing Date: 7 March 1995 (07.03.95) (30) Priority Data: 207,547 7 March 1994 (07.03.94) US 215,082 18 March 1994 (18.03.94) US (60) Parent Applications or Grants (63) Related by Continuation US 08/207,547 (CIP) Filed on 7 March 1994 (07.03.94) US 08/215,082 (CIP) Filed on 18 March 1994 (18.03.94) (71) Applicant (for all designated States except US): YALE UNIVERSITY [US/US]; 451 College Street, New Haven, CT 06511 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): YUAN, Yan [CN/US]; Apartment 2, 550 Whitney Avenue, New Haven, CT 06511 (US). GUERRIER-TAKADA, Cecilia [BR/US]; 57 Ramsdell Street, New Haven, CT 06515 (US). ALTMAN, Sidney [US/US]; 71 Blake Road, Hamden, CT 06517 (US).		(74) Agent: PABST, Patrea, L.; Arnall Golden & Gregory, 2800 One Atlantic Center, 1201 West Peachtree Street, Atlanta, GE 30309-3450 (US). (81) Designated States: AU, CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(54) Title: TARGETED CLEAVAGE OF RNA USING RIBONUCLEASE P TARGETING AND CLEAVAGE SEQUENCES			
(57) Abstract <p>It has been discovered that any RNA can be targeted for cleavage by RNase P from prokaryotic or eukaryotic cells using a suitably designed oligonucleotide ("external guide sequence", or EGS) to form a hybrid with the target RNA, thereby creating a substrate for cleavage by RNase P <i>in vitro</i>. The EGS hydrogen bonds to the targeted RNA to form a partial tRNA like structure including the aminoacyl acceptor stem, the T stem and loop, and part of the D stem. An EGS can be modified both by changes in sequence and by chemical modifications to the nucleotides. The EGS can be a separate molecule or can be combined with an RNase P catalytic RNA sequence to form a single oligonucleotide molecule ("RNase P internal guide sequence" or RIGS). Methods are also disclosed to randomly select and to express a suitable EGS or RIGS <i>in vivo</i> to make a selected RNA a target for cleavage by a host cell RNase P or introduced RIGS, thus preventing expression of the function of the target RNA. The methods and compositions should be useful to prevent the expression of disease- or disorder-causing genes <i>in vivo</i>.</p>			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhstan	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	Senegal
CN	China	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Viet Nam
GA	Gabon				

TARGETED CLEAVAGE OF RNA USING RIBONUCLEASE P TARGETING AND CLEAVAGE SEQUENCES

Background Of The Invention

5 This invention is in the general area of genetic engineering of nucleic acid sequences, especially chemically modified external guide sequences and catalytic RNA sequences linked to guide sequences.

 There are several classes of ribozymes now known which are involved in the cleavage and/or ligation of RNA chains. A ribozyme is
10 defined as an enzyme which is made of RNA, most of which work on RNA substrates. Ribozymes have been known since 1982, when Cech and colleagues (*Cell* 31: 147-157) showed that a ribosomal RNA precursor in *Tetrahymena*, a unicellular eukaryote, undergoes cleavage catalyzed by elements in the RNA sequence to be removed during the
15 conversion of the rRNA precursor into mature rRNA. Another class of ribozyme, discovered in 1983, was the first to be shown to work in trans, that is, to work under conditions where the ribozyme is built into one RNA chain while the substrate to be cleaved is a second, separate RNA chain. This ribozyme, called M1 RNA, was characterized in 1983 by
20 Altman and colleagues as responsible for the cleavage which forms mature 5' ends of all transfer RNAs (tRNAs) in *E. coli*. Analogous RNA-containing enzymes concerned with tRNA synthesis have since been found in all cells in which they have been sought, including a number of human cell lines, though the relevant eukaryotic RNAs have not yet been
25 shown to be catalytic by themselves *in vitro*.

 The discovery and characterization of this catalytic RNA is reviewed by Sidney Altman, in "Ribonuclease P: An Enzyme with a Catalytic RNA Subunit" in *Adv. Enzymol.* 62: 1-36 (1989). The activity was first isolated from *E. coli* extracts, and subsequently determined to be
30 a ribonucleoprotein having two components, an RNA component called M1 and a protein component called C5. The RNA cleaved substrates in a true enzymatic reaction, as measured using Michaelis-Menten kinetics. M1 was determined to be solely responsible for substrate recognition and

C5 was determined to alter k_{cat} but not K_M , as reported by Guerrier-Takada *et al.*, *Cell* 35: 849 (1983) and McClain *et al.*, *Science* 238: 527 (1987). Sequencing showed that M1 RNA is 377 nucleotides long, M_r approximately 125,000, and that the protein consists of 119 amino acids, M_r approximately 13,800, as reported by Hansen *et al.*, *Gene* 38: 535 (1987).

Cleavage of precursor tRNA molecules by the RNA component of eubacterial RNase P is described by Guerrier-Takada *et al.*, *Cell* 35, 849 (1983) and reviewed by Altman, *Adv. Enzymol.* 62: 1 (1989).

U.S. Patent No. 5,168,053 entitled "Cleavage Of Targeted RNA By RNase P" to Altman *et al.*, discloses that it is possible to target any RNA molecule for cleavage by bacterial RNase P by forming a nucleotide sequence part of which is complementary to a targeted site and which includes a terminal 3'-NCCA, wherein the sequence is designed to hybridize to the targeted RNA so that the bacterial RNase P cleaves the substrate at the hybrid base-paired region. Specificity is determined by the complementary sequence. The sequence is preferably ten to fifteen nucleotides in length and may contain non-complementary nucleotides to the extent this does not interfere with formation of several base pairs by the complementary sequence which is followed by NCCA at the 3' end.

As described in WO 92/03566 to Yale University, ribonuclease P (RNase P) from *E. coli* can cleave oligoribonucleotides that are found in hydrogen-bonded complexes that resemble the aminoacyl stem and include the 5' leader sequence of tRNA precursors, -NCAA. Human RNase P cannot cleave *in vitro* the 5' proximal oligoribonucleotide in the simple complexes cleaved by RNase P from *E. coli*, but can do so when the 3' proximal oligoribonucleotide is bound to an external guide sequence (EGS) to form a structure resembling portions of a tRNA molecule. The EGS can include a complementary sequence to a target substrate of at least eleven nucleotides, seven bases which hydrogen bond to the targeted sequence to form a structure akin to the

aminoacyl acceptor stem of a precursor tRNA, and four nucleotides which base pair with the targeted sequence to form a structure akin to the dihydroxyuracil stem. WO 92/03566 does not disclose EGS for prokaryotic RNase P with fewer than seven complementary nucleotides.

5 WO 93/22434 to Yale University discloses an EGS for human RNase P. As described in WO 93/22434, an EGS for human RNase P consists of a sequence which, when in a complex with the target substrate molecule, forms a secondary structure resembling that of a tRNA cloverleaf, or a substantial part of it, and that results in cleavage of the
10 target RNA by RNase P. The sequence of the EGS of WO 93/22434 is derived from any tRNA except that the D stem and aminoacyl stem are altered to be complementary to the target substrate sequence. WO 93/22434 also discloses EGS with either the anticodon loop and stem or the extra loop deleted, and EGS where the sequence of the T loop and
15 stem are changed. WO 93/22434 does not disclose eukaryotic EGS comprising only a region complementary to the target RNA and a region forming a structure similar to only the T stem and loop of tRNA. Neither WO 92/03566 nor WO 93/22434 discloses EGS having chemically modified nucleotides.

20 It is therefore an object of the present invention to provide methods and compositions for specifically cleaving targeted RNA sequences using linked catalytic RNA and minimal guide sequences.

It is another object of the present invention to provide chemically modified external guide sequences for RNase P with enhanced
25 resistance to nuclease degradation.

It is another object of the present invention to provide a method for selecting external guide sequences, and linked catalytic RNA and guide sequences, that cleave a target RNA with increased efficiency.

It is a further object of the present invention to provide
30 methods and compositions for specifically cleaving RNA, both *in vitro* and *in vivo* within eukaryotic cells, for the treatment of disease conditions which involve RNA transcription or translation, such as diseases caused

by RNA and DNA viruses and expression of excessive or pathogenic proteins from mRNA, or of excessive or pathogenic RNA, itself.

Summary Of The Invention

5 Any RNA can be targeted for cleavage by RNase P, using a suitably designed oligonucleotide ("external guide sequence") to form a hybrid with the target RNA, thereby creating a substrate targeted for cleavage by RNase P. The EGSs contain sequences which are complementary to the target RNA and which forms secondary and tertiary
10 structure akin to portions of a tRNA molecule. A eukaryotic EGS must contain at least seven nucleotides which base pair with the target sequence 3' to the intended cleavage site to form a structure like the amino acyl acceptor stem, nucleotides which base pair to form a stem and loop structure similar to the T stem and loop, followed by at least three
15 nucleotides that base pair with the target sequence to form a structure like the dihydroxyuracil stem. The EGS can be made more resistant to nuclease degradation by including chemically modified nucleotides or nucleotide linkages.

The external guide sequence and the RNase P catalytic RNA
20 can be used together as separate molecules. Alternatively, the two sequences can be combined into a single oligonucleotide molecule possessing both targeting and catalytic functions. Such a combined oligonucleotide, termed an RNase P internal guide sequence (RIGS), increases the kinetic efficiency of cleavage by reducing the number of
25 reactants and by keeping the targeting and catalytic elements in close proximity. Chemically modifying the nucleotides and phosphate linkages of EGS molecules and RIGS molecules make the oligonucleotides more resistant to nuclease degradation.

Methods are also disclosed to select RIGS molecules and EGS
30 molecules having increased substrate affinity and utility *in vivo* to cleave or target cleavage of a selected RNA, thus preventing expression of the function of the target RNA. The methods and compositions should be

useful to prevent the expression of disease- or disorder-causing genes *in vivo*.

As described in the examples, an RIGS was constructed by linking a guide sequence to M1 RNA (M1GS RNA). M1GS RNA can act
5 as a sequence-specific endonuclease and can cleave target RNAs that base pair with the guide sequence just as group I introns do. A custom-designed M1GS RNA cleaves the mRNA that encodes thymidine kinase (TK) of human herpes simplex virus 1 (HSV-1) *in vitro*. When this
10 M1GS RNA is expressed in mammalian cells in tissue culture, it reduces the level of expression of TK by decreasing the amount of the target TK mRNA.

Brief Description Of The Drawings

Figure 1 is the proposed secondary structure of the chimeric
15 substrate used in the selection procedure (Sequence ID No. 1). The italicized sequence is from CAT mRNA and the remaining sequence, aside from changes made to assure hydrogen bonding specifically to CAT mRNA, is based on the sequence of *E. coli* tRNA^{Tyr}. The nine
20 nucleotides that were randomized are each indicated by N. Some of the EGS analogs for various parts of a tRNA are also depicted: acceptor stem, T loop, anticodon loop, and D-loop.

Figure 2 depicts the general scheme for *in vitro* selection and amplification of chimeric substrates with enhanced efficiency for directing
25 human RNase P cleavage. "Reverse transcriptase" denotes that the DNA polymerase capability of reverse transcriptase was used to create double-stranded template DNA from the overlapping DNA oligonucleotides SEC-1A (Sequence ID No. 2) and SEC-1B (Sequence ID No. 3). "RNA-PCR" refers to reverse transcription coupled to PCR.

Figure 3 is the sequence and secondary structures of the
30 precursor to tRNA^{Tyr} (Sequence ID No. 4) and a complex of a substrate and EGSΔ(1-18) (Sequence ID No. 5), which is derived from tRNA^{Tyr} but lacks the first eighteen nucleotides from the 5' end of the mature tRNA^{Tyr}.

pTyr: *E. coli* tRNA^{Tyr} precursor; pAva: substrate (target) RNA with 5' leader sequence and the first fourteen nucleotides of *E. coli* tRNA^{Tyr} (Sequence ID No. 6).

Figures 4A, 4B, 4C, 4D, 4E, and 4F are the proposed secondary structures of complexes of CAT mRNA and various EGS. 4A and 4D depict complexes between the complementary region CAT mRNA (Sequence ID No. 7) and EGS^{CAT} (Sequence ID No. 8). 4B and 4E are complexes of CAT mRNA (Sequence ID No. 7) and EGS 9 (Sequence ID No. 9). 4C and 4F are complexes of CAT mRNA (Sequence ID No. 7) and EGS^{CAT}ΔAC (Sequence ID No. 10). Hollow arrows denote sites of cleavage by human RNase P. In Figures 4A, 4B and 4C solid arrowheads denote the sites of cleavage by RNase T1, and arrows denote sites of cleavage by RNase T2. In Figures 4D, 4E and 4F, sites of cleavage by cobra venom nuclease are indicated by solid arrows.

Figure 5 shows rates (nmol/min) of human RNase P cleavage of CAT mRNA directed by twelve individual EGS RNAs. Nine individual EGS RNAs: EGS 1 (Sequence ID No. 17), EGS 4 (Sequence ID No. 19), EGS 5 (Sequence ID No. 20), EGS 6 (Sequence ID No. 21), EGS 8 (Sequence ID No. 22), EGS 9 (Sequence ID No. 9), EGS 12 (Sequence ID No. 26), EGS 14 (Sequence ID No. 27), EGS 18 (Sequence ID No. 31), were prepared by *in vitro* selection. EGS 19 and EGS^{CAT}ΔAC (Sequence ID No. 10) were prepared by *in vitro* mutagenesis. Results are presented as initial rates (nmol/mol) of cleavage of substrate by RNase P during the linear phase of each reaction.

Figures 6A, 6B, and 6C are sequence and proposed secondary structures of EGS for herpes simplex virus thymidine kinase mRNA (Sequence ID No. 35). Figure 6A is an EGS (Sequence ID No. 36) forming an aminoacyl acceptor stem, T loop and stem, variable loop and stem, anticodon loop and stem, and D stem. Figure 6B is an EGS (Sequence ID No. 37) wherein a G in the T loop is substituted for a C in the EGS of Figure 6A. Figure 6C is an EGS (Sequence ID No. 38) where the anticodon loop and stem of the EGS of Figure 6A is deleted.

Figure 7 is a graph of percent TK mRNA expression for cell lines: CL-CAT, CL-109, CL-104, and CL-112.

Figure 8 is the sequence and structure of H1 RNA (Sequence ID No. 11) from Altman *et al.*, *Genomics* 18: 418-422 (1993).

5 Figure 9 is a schematic representation of complexes that are formed by M1GS RNA (Sequence ID No. 51 and Sequence ID No. 52) and two substrates (Sequence ID No. 53 and Sequence ID No. 54).

 Figure 10 is a diagram of TK and CaT RNA substrate RNAs. The two 13-nt segments highlighted in the TK RNA and CAT RNA
10 sequences (Sequence ID No. 55 and Sequence ID No. 56) are complementary to the guide sequences in M1TK13 and M1CAT13 RNAs, respectively. The cleavage sites are indicated by arrows.

 Figure 11 is a schematic representation of the tk46 substrate sequence (Sequence ID No. 57) and the structure of RNA enzymes used
15 (Sequence ID No. 58, Sequence ID No. 59, Sequence ID No. 60, Sequence ID No. 61, Sequence ID No. 62, Sequence ID No. 63, Sequence ID No. 64, and Sequence ID No. 65). The arrow marks the expected site of cleavage. The sequences shown in [the M1GS RNAs] bold type are the guide sequences that contain the 3' CCA sequence and
20 the sequence complementary to the TK sequence.

Figure 12 is a schematic representation of retroviral vectors (Δ M1TK, M1TK, and NB2) containing M1GS RNAs or an EGS sequence complementary to the influenza viral protein PB2.

 Figure 13 is a graph of levels of expression (percent of control
25 expression) of TK mRNA and protein in ψ CRE cells and in other cell lines. Values are averages of results from four independent experiments. Results from the four experiments varied within 5% in absolute terms. Solid bars: TK mRNA; open bars; TK protein.

Detailed Description Of The Invention

Guide Sequence

As used herein, a guide sequence (GS) is an oligonucleotide that targets a substrate RNA molecule for cleavage by a catalytic RNA
5 having the activity of RNase P catalytic RNA. A guide sequence may be a separate molecule, termed an external guide sequence, or combined in a single molecule with catalytic RNA. Such a combined molecule is referred to herein as an RNase P internal guide sequence (RIGS).

A. Eukaryotic RNase P Targeting Sequence.

10 A guide sequence for human RNase P consists of a sequence which, when in a complex with the target substrate molecule, forms a secondary structure resembling that of a tRNA cloverleaf or a part of it. As used herein, the term "resembling a precursor tRNA" means a complex formed by the GS with target RNA substrate to resemble a
15 sufficient portion of the tRNA secondary and tertiary structure to result in cleavage of the target RNA by RNase P. The sequence of the GS can be derived from any tRNA except that the D stem and aminoacyl stem have to be altered to be complementary to the target substrate sequence. These altered stems are referred to as recognition arms. The recognition arm
20 corresponding to the aminoacyl stem is referred to as the A recognition arm and the recognition arm corresponding to the D stem is referred to as the D recognition arm. The remaining portion of the guide sequence, which is required to cause RNase P catalytic RNA to interact with the GS/target sequence complex, is herein referred to as RNase P binding
25 sequence. The presence of a 3'-CCA on an EGS enhances the efficiency of *in vitro* reaction with the human RNase P by about 35%. The anticodon loop and stem and extra loop can separately be deleted and the sequence of the T loop and stem can be changed without decreasing the usefulness of the guide sequence and, in the case of the anticodon stem
30 and loop deletion, increases the efficiency of the reaction by about ten fold. Changes in other parts of an EGS can increase its efficiency about one hundred fold.

The desired secondary structure is determined using conventional Watson-Crick base pairing schemes to form a structure resembling a tRNA, that is, having structure as described below. The specific sequence of the hydrogen bonded regions is not as critical, as long as the desired structure is formed. All tRNAs, including tRNAs from a wide variety of bacteria and eukaryotes, conform to the same general secondary structure. This is typically written in the form of a cloverleaf, maintained by hydrogen-bonded base pairing between short complementary regions. The four major arms are named for their structure or function: The acceptor arm consists of a 3' terminal CCA_{OH} plus a variable fourth nucleotide extending beyond the stem formed by base-pairing the 5' and 3' segments of the molecule. The other arms consist of base-paired stems and unpaired loops. The "T" arm is named for the presence of the ribothymidine nucleotide and contains seven unpaired bases in the loop. The anticodon arm always contains the anticodon triplet in the center of the loop and consists of seven unpaired bases. The D arm is named for the presence of the base dihydrouridine in the loop, another of the chemically modified bases in tRNA, and includes between eight and twelve unpaired bases. Positions are numbered from 5' to 3' according to the most common tRNA structure, which has 76 residues. The overall range of tRNA lengths is from 74 to 95 bases. The variation in length is caused by differences in the structure of two of the arms, the D arm and the extra or variable arm, which lies between the T and anticodon arms, which can contain between three and five bases, or between 13 and 21 bases with a stem of about five bases. The base pairing that maintains the secondary structure is virtually invariant: there are always seven base pairs in the acceptor stem, five in the T arm, five in the anticodon arm, and three or four in the D arm.

As used herein, a hybrid structure, consisting of an EGS hydrogen bonded to an RNA substrate, having secondary structure resembling a precursor tRNA under conditions promoting cleavage by RNase P of the substrate at the nucleotide at the 5' end of the base-paired

region, preferably includes a D stem, an aminoacyl stem, and a T loop and stem, where the sequence of the latter may be changed compared to the sequence and detailed structure found in the parent molecule.

A few nucleotides are always found in the same positions in
5 90 to 95% of tRNAs, with some additional nucleotides being semiconserved or semivariant. This is not an absolute requirement in the GS, as long as the sequence is complementary to the target and forms the secondary structure characteristic of the tRNA. In fact, the sequence forming the aminoacyl stem and D loop and stem are changed in the GS
10 to be complementary to the target RNA.

The base paired double-helical stems of the secondary structure are maintained in the tertiary structure, creating two double helices at right angles to each other. The acceptor stem and the T stem form one continuous double helix with a single gap; the D stem and the
15 anticodon stem form another continuous double helix, also with a gap. Many of the invariant and semi-invariant bases are involved in the tertiary structure.

The complementary sequences will generally consist of eleven nucleotides, or, under certain conditions may consist of as few as seven
20 nucleotides, in two blocks which base pair with the target sequence and which are separated by two unpaired nucleotides in the target sequence, preferably UU, wherein the two blocks are complementary to a sequence 3' to the site targeted for cleavage.

B. Prokaryotic RNase P Targeting Sequence.

25 The requirements for a GS functional with prokaryotic RNase P are less stringent than those for a eukaryotic GS. The critical elements of a prokaryotic GS are (1) nucleotide sequence which specifically binds to the targeted RNA substrate to produce a short sequence of base pairs 3' to the cleavage site on the substrate RNA and (2) a terminal 3'-NCCA,
30 where N is any nucleotide, preferably a purine. The sequence generally has no fewer than four, and more usually six to fifteen, nucleotides complementary to the targeted RNA. It is not critical that all nucleotides

be complementary, although the efficiency of the reaction will vary with the degree of complementarity. The rate of cleavage is dependent on the RNase P, the secondary structure of the hybrid substrate, which includes the targeted RNA and the presence of the 3'-NCCA in the hybrid
5 substrate.

Ribonuclease P

Ribonuclease P is an enzyme consisting of protein and RNA subunits that cleaves tRNA precursors to generate the 5' termini of
10 tRNAs. This essential enzymatic activity has been found in all cell types examined, both prokaryotic and eukaryotic. During the studies on recognition of substrate by RNase P, it was found that *E. coli* RNase P can cleave synthetic tRNA-related substrates that lack certain domains, specifically, the D, T and anticodon stems and loops, of the normal tRNA
15 structure. A half-turn of an RNA helix and a 3' proximal CCA sequence contain sufficient recognition elements to allow the reaction to proceed. The 5' proximal sequence of the RNA helix does not have to be covalently linked to 3' proximal sequence of the helix. The 3' proximal sequence of the stem can be regarded as a "guide sequence" because it
20 identifies the site of cleavage in the 5' proximal region through a base-paired region.

RNase P from *E. coli* and human cells have similar but not identical biochemical properties. Their RNA components have similar secondary structures. However, the substrate range of human RNase P is
25 much narrower than that of the *E. coli* enzyme. For example, although *E. coli* RNase P can cleave a synthetic tRNA-related substrate that lacks three specific domains of the normal tRNA structure, the human enzyme and the structurally similar enzyme from the yeast, *S. cerevisiae*, cannot cleave the same substrate. However, the *E. coli* RNase P can cleave a
30 synthetic tRNA-related substrate that is also cleaved by the human RNase P. Altman *et al.*, *Genomics* 18: 419-422 (1993), describes several

mammalian RNase P catalytic RNAs and identifies common features and differences.

As used herein, unless otherwise specified, RNase P refers to the RNase P in the cell in which the RNA to be cleaved is located, whether endogenous, added to the cell, or as used *in vitro*. Many of the techniques described herein are known to those skilled in the art, as are methods for making, and sources of, reagents. The teachings of any references cited herein with respect to methods and reagents are specifically incorporated herein, as well as for the purpose of demonstrating the scope and level of skill in the art.

It is not necessary to provide RNase P activity if the cleavage is to occur in bacterial cells or intracellularly in the nucleus since all eukaryotic cells contain RNase P in their nuclei. RNase P must be supplied if cleavage is to occur in the cytoplasm of eukaryotic cells. As used herein for ease of convenience, RNase P refers to the ribonucleoprotein consisting of prokaryotic or eukaryotic analogues of the *E. coli* C5 protein and M1 RNA, regardless of source, whether isolated, or produced by chemical synthesis. The RNA subunit of RNase P also can be transcribed from a gene. The eukaryotic RNase P RNA subunit is referred to as H1 RNA. The RNA subunit need not necessarily manifest catalytic activity in the absence of protein subunits *in vitro*.

A. Endogenous RNase P.

The sequence and proposed secondary structure of H1 RNA, the RNA component of human RNase P, was reported by Altman *et al.* (1993), the teachings of which are generally known in the art. The sequence and proposed structure of H1 RNA is shown in Figure 8 (Sequence ID No. 11). The sequence and proposed secondary structure of M1 RNA, the RNA component of *E. coli* RNase P, was reported by James *et al.*, *Cell* 52: 19 (1988), the teachings of which are generally known in the art. The sequence of M1 RNA is included as Sequence ID No. 40.

Because of the similarity in secondary structure and substrate specificity among the RNase P's of diverse origin, it is possible to use an EGS designed to maximize efficiency of cleavage for the RNase P in question using techniques described herein to target any RNA in any cell, even though the catalytically active RNA subunits may have distinctly different sequences. See Altman, *Ann. Rev. Enzymology* 62: 1-39 (1989); Altman, *J. Biol. Chem.* 265: 20053-20056 (1990). Secondary structure is defined by intramolecular associations of complementary sequences. Base pairs can be canonical, A/U and G/C, or non-canonical, G/U, A/G, etc.

B. Exogenous RNA Having Catalytic Activity.

An EGS can also be used in combination with an RNA sequence that demonstrates enzymatic activity in the presence or absence of a protein. That RNA sequence can be represented by a molecule like the entire H1 or M1 RNA molecule, any functionally equivalent molecule of prokaryotic or eukaryotic origin or derivation, or any portion thereof shown to have catalytic activity, either alone or in combination with a protein. Such a catalytic RNA is referred to herein as RNase P catalytic RNA and its sequence is referred to as an RNase P catalytic sequence. An RNA as described above is considered an RNase P catalytic RNA regardless of source, whether isolated, produced by chemical synthesis, or transcribed from a gene. As noted above, an EGS effective to convert a targeted sequence into a substrate for human RNase P, will also be effective in making the substrate a target for procaryotic RNase P.

RNase P catalytic RNA can be derived from naturally occurring RNase P catalytic RNAs, for example, by deleting portions and by making nucleotide substitutions. Such derived catalytic RNAs need only retain enough of the catalytic activity of naturally occurring RNase P catalytic RNA to cleave target RNA. A preferred method of generating RNase P catalytic sequences is by *in vitro* evolution as described below.

There are two principle situations in which RNase P catalytic RNA or RNase P is utilized in combination with EGS: *in vitro* in the absence of cells or cellular RNase P and in circumstances wherein the

RNA to be cleaved is located in a portion of a cell not containing endogenous RNase P. In the latter case, the genes encoding the analogs of M1 RNA and C5 protein, as defined above, or the human or other eukaryotic equivalents thereof, are introduced into the cell at the desired
5 location for cleavage using a suitable vector or other method known to those skilled in the art for introduction and expression of a gene in a cell.

RNase P Internal Guide Sequences

A guide sequence and the catalytic RNA subunit of an RNase
10 P can be linked to form a single oligonucleotide molecule possessing both the targeting function of an EGS and cleavage function of RNase P catalytic RNA. Such a combination, in a single oligonucleotide molecule, is referred to as an RNase P internal guide sequence (RIGS). An RIGS can be used to cleave a target RNA molecule in the same manner as EGS.

15 RIGSs can be formed by linking a guide sequence to an RNase P catalytic sequence by any suitable means. For example, an EGS and RNase P catalytic RNA can be prepared as separate molecules which are then covalently linked *in vitro*. Alternatively, a complete RIGS can be synthesized as a single molecule, either by chemical synthesis, or by
20 *in vitro* or *in vivo* transcription of a DNA molecule encoding linked GS and RNase P catalytic sequence. The linkage between the GS and RNase P domains of an RIGS can have any form that allows the domains to cleave a target RNA. For example, the two domains could be joined by an oligonucleotide linker. Preferably, the linker will be composed of an
25 ordinary nucleotides joined by phosphodiester bonds. The GS and RNase P catalytic sequence components can be joined in either order, with the RNase P catalytic sequence linked to either the 3' end or 5' end of the GS component.

RIGSs can be used for cleavage of target RNA both *in vitro*
30 and *in vivo*. *In vitro*, the RIGS can function without RNase P protein components *in vitro*, although activity of the RIGS can be increased by the addition of RNase P protein components. *In vivo*, endogenous RNase

P proteins will stimulate activity of the RIGS. The activity of both prokaryotic- and eukaryotic-based RIGSs are expected to be enhanced by the presence of either prokaryotic or eukaryotic RNase P protein components.

5

Method For Producing EGSs And RIGSs Having Enhanced Efficacy

EGSs and RIGSs having enhanced binding affinity as measured by decreased energy of binding can be designed by *in vitro* evolution. Such a method can be used to identify RNA molecules with desired properties from pools of molecules that contain randomized sequences. As demonstrated more clearly in the examples, appropriately modified, these methods can be used for the isolation of efficient EGSs and RIGSs. These new EGSs, when complexed with an exemplary target RNA, CAT mRNA substrate (Sequence ID No. 7), allow cleavage of the target by human RNase P at rates similar to those achieved with natural substrate.

The general selection scheme is depicted in Figure 2. In each round of selection, the pool of RNAs is digested with human RNase P, or with the RIGS, and the cleaved products are isolated by electrophoresis and then amplified to produce progeny RNAs. One of the template-creating oligonucleotides is used as the 5' primer for the polymerase chain reaction (PCR) in order to allow restoration of the promoter sequence and the leader sequence of the chimeric RNA for the next cycle of selection. The stringency of selection is increased at each cycle by reducing the amount of enzyme and the time allowed for the cleavage reaction, such that only those substrates that are cleaved rapidly by the enzyme are selected.

In the first three rounds of selection, RNA substrates are digested with an appropriate amount of human RNase P, for example, 3.6 units, or the equivalent activity of the RIGS. One unit of human RNase P is defined as that amount of enzyme that cleaves 1 pmol of precursor to

tRNA^{Tyr} from *E. coli* in 30 min at 37°C. For assays in subsequent rounds of selection, the amount of enzyme is reduced, and the incubation time is shortened so that less than 20 percent of the substrate is cleaved. Cleavage products are separated from uncleaved substrates by electrophoresis and RNA extracted.

The purified cleavage product RNAs are reverse transcribed and amplified by PCR. The double-stranded DNA generated by PCR regains the promoter sequence and the leader sequence from the sequence in the primer, and is then used as a template for transcription of RNA for the next round of selection. After eight cycles of selection, the resulting pool of double-stranded DNAs is cloned into an appropriate vector and sequenced.

In order to test the abilities of EGSs or RIGSs derived from the individual variants, sequences corresponding to the GS segment of each chimeric tRNA are amplified by PCR, and RNAs transcribed with an appropriate RNA polymerase. RNA cleavage by the selected EGS or RIGS is then assayed. Sequences in common in the most active EGS and RIGSs are then determined and new EGS and RIGSs designed.

As described in the examples, simulation of evolution *in vitro* was used to select EGSs that bind strongly to a target substrate mRNA and that increase the efficiency of cleavage of the target by human ribonuclease P to a level equal to that achieved with natural substrates. The most efficient EGSs from tRNA precursor-like structures with the target RNA, in which the analog of the anticodon stem has been disrupted, an indication that selection for the optimal substrate for ribonuclease P yields an RNA structure different from that of present-day tRNA precursors.

Chemically Modified EGS And RIGS

Although chemically unmodified oligoribonucleotides can function as effective EGS or RIGS in a nuclease-free environment, the shorter half-life in serum and inside cells reduces their effectiveness as

therapeutics. Chemical modifications can be made which greatly enhance the nuclease resistance of EGS and RIGS without compromising their biological function of inducing or catalyzing cleavage of RNA target. For example, one or more of the bases of an EGS or RIGS construct can be replaced by 2' methoxy ribonucleotides or phosphorothioate deoxyribonucleotides using available nucleic acid synthesis methods. Synthesis methods are described by, for example, Offensperger *et al.*, *EMBO J.* 12: 1257-1262 (1993); PCT WO 93/01286 by Rosenberg *et al.* (synthesis of sulfurthioate oligonucleotides); Agrawal *et al.*, *Proc. Natl. Acad. Sci. USA* 85: 7079-7083 (1988); Sarin *et al.*, *Proc. Natl. Acad. Sci. USA* 85: 7448-7794 (1989); Shaw *et al.*, *Nucleic Acids Res* 19: 747-750 (1991) (synthesis of 3' exonuclease-resistant oligonucleotides containing 3' terminal phosphoroamidate modifications); all of which are hereby incorporated herein by reference.

It is well documented in the current literature that degradation of oligonucleotide analogues is mainly attributable to 3'-exonucleases. Several studies have also demonstrated that various 3'-modifications can greatly decrease the nuclease susceptibility of these analogues. Thus, another method to reduce susceptibility to 3' exonucleases is introduction of a free amine to a 3' terminal hydroxyl group of the EGS or RIGS molecule, as described, for example, by Orson *et al.*, *Nucl. Acids Res.* 19: 3435-3441 (1991). Furthermore, cytosines that may be present in the sequence can be methylated, or an intercalating agent, such as an acridine derivative, can be covalently attached to a 5' terminal phosphate to reduce the susceptibility of a nucleic acid molecule to intracellular nucleases. Examples of this are described in Maher *et al.*, *Science* 245: 725-730 (1989) and Grigoriev *et al.*, *J. Biol. Chem.* 267: 3389-3395 (1992).

Another class of chemical modifications is modification of the 2' OH group of a nucleotide's ribose moiety, which has been shown to be critical for the activity of the various intracellular and extracellular nucleases. Typical 2' modifications are the synthesis of 2'-O-Methyl oligonucleotides, as described by Paoella *et al.*, *EMBO J.* 11: 1913-1919

(1992), and 2'- fluoro and 2'-amino-oligonucleotides, as described by Picken *et al.*, *Science* 253: 314-317 (1991), and Heidenreich and Eckstain, *J. Biol. Chem* 267: 1904-1909 (1992). Portions of EGS and RIGS molecules can also contain deoxyribonucleotides. Such substitutions
5 improve nuclease resistance by eliminating the critical 2' OH group.

Application Of EGS And RIGSs As Laboratory Or Clinical Reagents

External guide sequences and RIGSs have applications as *in*
10 *vitro* reagents, in a similar fashion to restriction enzymes, and as therapeutic agents, for cleavage and inactivation of specific host cell RNA or RNA coded for by pathogenic organisms such as bacteria or viruses, as demonstrated by the following examples.

As used herein, an EGS or RIGS is an oligonucleotide
15 molecule. It is to be understood, however, that for therapeutic purposes, a DNA molecule encoding an EGS molecule or encoding an RIGS molecule could be utilized. Accordingly, unless otherwise specified, administration of an EGS or RIGS encompasses both the RNA molecule that hydrogen bonds to a target nucleic acid sequence that is cleaved, as
20 well as a DNA molecule encoding the RNA molecule, which is expressed under conditions wherein the RNA molecule functions as an EGS or RIGS.

An external guide sequence can be brought into contact with any RNA having sequences complementary to the EGS, in the presence of
25 RNase P, and the RNA will be cleaved at the targeted site. In this manner, the activity of endogenous RNase P in any cell, such as the RNase P of human cells, can be directed to destroy specific messenger, viral or other RNAs by the use of an appropriate EGS RNA.

A. Reagents For *In vitro* Applications.

30 DNA restriction endonucleases are invaluable reagents for the molecular biologist. Patterns of restriction fragment sizes are used to establish sequence relationships between DNA molecules, and large

DNAs can be cleaved to give fragments of sizes useful for genetic engineering, sequencing, and studying protein binding. RNA processing enzymes can be utilized under conditions such that they also cleave RNA with considerable sequence specificity.

- 5 Specific ribozymes can be prepared by combining the specific guide sequence with RNase P or functional equivalents thereof. In the preferred embodiment, the external guide sequence and the RNase P catalytic RNA are separate; alternatively, the two sequences can be joined either directly or via a linker. The linker can be any molecule that can be
- 10 covalently bound to oligonucleotides. Numerous linkers are known to those of skill in the art. A preferred linker is an oligonucleotide because it allows direct synthesis of the complete RIGS.

B. Therapeutics.

15 1. **Determination and Preparation of Complementary Sequences.**

- Any cellular gene product expressed as RNA, including proteins encoded by mRNA and structural RNAs themselves, can be targeted for inactivation by RNase P, or directly by an RIGS using sequences engineered to include appropriate regions of sequence and/or
- 20 structure for binding to the targeted RNA and the desired site of cleavage. The cellular gene product could be a product of an oncogene with an altered sequence, such as the *ras* gene product; where the product is not a normal cell component, a viral protein, such as one encoded by an essential gene for HIV replication; or a bacterial protein.

- 25 In many cases, the critical genes an infective or pathological agent have been isolated and sequenced. Appropriate complementary sequences can be synthesized using standard techniques, reagents, and equipment based on these known sequences.

2. Preparation of an appropriate pharmaceutical composition for delivery of the EGS or RIGS to the targeted RNA.

There are two primary mechanisms for delivering the EGS or
5 RIGS to intracellular RNA that has been targeted for cleavage: diffusion
and via a vector.

As discussed above, any RNA that is important in a disease
process can be targeted and appropriate complementary sequences made
synthetically or by copying cloned sequence. For example, cancer
10 regulatory genes can be targeted. Since RNase P is predominantly found
in the nucleus of eukaryotic cells, the infectious agents most likely to be
inhibited by administration of appropriate EGS to the infected cells are
those in which critical RNA sequences are transcribed in the nucleus.
Important examples of the viral agents that replicate in the cell nucleus
15 include herpes viruses, including herpes simplex virus, varicella-herpes
zoster virus, cytomegalovirus, and Epstein-Barr virus; hepatitis B virus;
adenoviruses; paramyxoviruses such as measles; and the retroviruses,
such as human immunodeficiency virus, HIV I, HIV II, HIV III and
HTLV-1. RIGSs should cleave target RNA in any area of the cell since
20 the catalytic activity is self-contained.

3. Vector-mediated delivery of EGS and RIGSs.

Preferred vectors are viral vectors such as the retroviruses
which introduce EGS and RIGS directly into the nucleus where it is
transcribed and released into the nucleus. Under the appropriate
25 conditions, the EGS or RIGS will hybridize to the targeted RNA and the
endogenous RNase P or RIGS will cleave the hybridized RNA at the 5'
side of the hybrid region.

Methods for using retroviral vectors for gene therapy are
described in U.S. Patent Nos. 4,868,116 and 4,980,286, and PCT
30 application PCT/US89/03794 and PCT/US89/00422, the teachings of
which are incorporated herein.

Defective retroviral vectors, which incorporate their own RNA
sequence in the form of DNA into the host chromosome, can be

engineered to incorporate EGS and RIGSs into the host, where copies will be made and released into the cytoplasm to interact with the target nucleotide sequences.

EGS and RIGSs are particularly useful as a means of targeted
5 therapy into hematopoietic cells of patients suffering from virus-induced disease of those cells, such as AIDS. The most efficacious methodology presently available for the introduction of specific genetic sequences into human cells involves the use of RNA-containing retroviruses which serve as vehicles or vectors for high efficiency gene transfer into human cells.

10 RNase P-based therapy can also be used as a means of preventing the spread of HIV-1 and or providing a HIV-1 resistant population of T-cells that will be able to confer immune function to infected individuals. Patients who have been recently diagnosed as having antibodies to HIV-1, but who do not yet show AIDS symptomatology, are
15 the preferred candidates for therapy. This procedure necessitates removal of some of the patient's bone marrow stem cells and subsequent partial cytoblation. The removed cells can be treated in the laboratory with appropriate EGS or RIGS compositions, using appropriate viral vectors, such as defective viral vectors, and then restored to the same individual.
20 The treated cells will develop in the patient into mature hematopoietic cells, including T-cells. These T-cells will have normal immune function and, most importantly, will be intracellularly immunized to prevent their destruction by any HIV-1 still present in the patient.

Bone marrow stem cells and hematopoietic cells are relatively
25 easily removed and replaced and provide a self-regenerating population of cells for the propagation of transferred genes. RNase P-based therapeutics allow the selective inactivation of other unwanted genes in cells, such as activated oncogenes, involved in the causation and maintenance of cancer cells.

30 In contrast to the approaches presently in use which are aimed at preventing or limiting infection with HIV, it should be possible to use RNase P-based technology to treat, and possibly to cure, HIV infection,

and related diseases of white blood cells which are subject to transformation by retroviral vectors carrying EGS or RIGS. Particular examples of diseases that may be treated using EGS and RIGS include not only HTLV-1, but also various retroviral-induced leukemias resulting from chromosomal translocations that produce chimeric RNAs which produce proteins that are unique to those cells and that can act as growth stimulators or oncogenes. Other types of transformed tissues that might be treatable include all cancer cells carrying identified oncogenes of known sequence.

10 **4. Topical and other EGS and RIGS compositions for local administration.**

EGS and RIGS may also be administered topically, locally or systemically in a suitable pharmaceutical carrier. *Remington's Pharmaceutical Sciences*, 15th Edition by E.W. Martin (Mark Publishing Company, 1975), the teachings of which are incorporated herein by reference, discloses typical carriers and methods of preparation. EGS and RIGS may also be encapsulated in suitable biocompatible microcapsules, microparticles or microspheres formed of biodegradable or non-biodegradable polymers or proteins or liposomes for targeting to phagocytic cells. Such systems are well known to those skilled in the art and may be optimized for use with the appropriate EGS and RIGSs.

Therapeutically the oligoribonucleotides are administered as a pharmaceutical composition consisting of an effective amount of the EGS or RIGS to inhibit transcription of a targeted RNA and a pharmaceutically acceptable carrier. Examples of typical pharmaceutical carriers, used alone or in combination, include one or more solid, semi-solid, or liquid diluents, fillers and formulation adjuvants which are non-toxic, inert and pharmaceutically acceptable. Such pharmaceutical compositions are preferable in dosage unit form, that is, physically discreet units containing a predetermined amount of the drug corresponding to a fraction or multiple of the dose which is calculated to produce the desired therapeutic response. It is essential that the oligonucleotides be delivered in a form

which prevents degradation of all of the oligonucleotide before it reaches the intended target site.

A preferred embodiment is an EGS or an RIGS administered as a viral vector, encoding the EGS or RIGS, or in a liposome, such that
5 an effective amount of EGS or RIGS is delivered. Generally these will produce a concentration between 1 μ M and 1 mM at the site of the cells to be treated. Such compounds and compositions can be formulated as a topical composition, for example, for application to a viral lesion such as that produced by herpes simplex virus. These will generally contain
10 between 1 μ M and 1 mM oligonucleotide per unit of carrier, or produce a concentration between 1 μ M and 1 mM at the site of the cells to be treated. Oral compositions, although not preferred, are in the form of tablets or capsules and may contain conventional excipients. Another preferred composition is a polymeric material applied locally for release
15 of EGS or RIGS. Still another preferred composition is a solution or suspension of the EGS or RIGS in an appropriate vector in combination with conventional pharmaceutical vehicles are employed for parenteral compositions, such as an aqueous solution for intravenous injection or an oily suspension for intramuscular injection.

20 For clinical applications, the dosage and the dosage regimen in each case should be carefully adjusted, utilizing sound professional judgment and consideration of the age, weight and condition of the recipient, the route of administration and the nature and gravity of the illness.

25 The present invention, EGS and RIGSs, will be further understood by reference to the following non-limiting examples.

Example 1: Cleavage of tRNA precursor fragments by human RNase P in the presence of EGS.

30 An EGS that can target RNA for cleavage by human RNase P was prepared using a small RNA fragment, pAva (Sequence ID No. 6, Figure 3) which contains a 5' precursor sequence and the first fourteen nucleotides from the 5' terminus of a tRNA. The leader sequence of *E*.

coli tRNA^{Tyr} precursor can be cleaved correctly when another piece of RNA, for example, EGSΔ1-18 (Sequence ID No. 5), which lacks the first eighteen nucleotides of the 5' terminus of mature tRNA^{Tyr} but retains the remaining 3' proximal sequence, is hybridized to the target RNA.

5 Human RNase P was partially purified from HeLa cells using the method of Bartkiewicz *et al.*, *Genes and Development* 3: 488-499 (1989). The substrates were prepared by *in vitro* transcription in the presence of [α -³²P]GTP. [α -³²P]GTP labelled pAva (Sequence ID No. 6) I RNA (28 nt) was mixed with unlabelled EGS RNA and the mixture was
10 incubated at 37°C in 50 mM Tris-Cl (pH 7.5), 100 mM NH₄Cl, and 10 mM MgCl₂ with enzyme for 30 min. Labelled pAva (Sequence ID No. 6) I RNA alone was also incubated with or without enzyme. Analysis by gel electrophoresis shows that the EGS plus enzyme resulted in cleavage of the RNA.

15 The 3' proximal oligonucleotide is the external guide sequence. Because the lengths of the leaders and their sequences, as well as the sequences of the mature domain, are not conserved among different precursor-tRNAs, the main determinants for human RNase P cleavage must be in some of the conserved structural features of various tRNAs.
20 This general idea is borne out by the fact that several other EGSs that did not mimic exactly the structure of parts of a tRNA did not target complementary RNAs. Examples of such changes include changes in the number of possible base pairs in the D or amino acyl stems, changes in positions 8 and 9 of the mature tRNA sequence, and a change from
25 cytosine to uracil at position 57. However, an EGS that lacked the anticodon stem and loop, or the variable stem and loop, resulted in efficient cleavage, indicating that these parts of the EGS, separately, were not essential for recognition of the target complex by the enzyme.

Accordingly, if an mRNA, rather than part of a precursor
30 tRNA sequence, is incorporated into the double-stranded stem region of a putative target complex, and the resulting hybrid contains the structural

features required of a substrate for human RNase P activity, the mRNA will be cleaved by human RNase P.

5 **Example 2: Specific cleavage of CAT mRNA *in vitro* by human RNase P using an external guide sequence.**

Examples provided below show the efficiency of cleavage of the mRNA for chloramphenicol acetyltransferase (CAT) by human RNase P. The sequence of the 5' oligoribonucleotide, as well as that of the EGS, depends on the choice of target site in the mRNA. The presence of
10 the appropriately designed EGS efficiently reduces CAT enzymatic activity *in vivo*, as well as promotes cleavage of CAT mRNA *in vitro*, indicating that this method should be of general use for gene inactivation.

An EGS custom-designed for the mRNA for chloramphenicol transacetylase (CAT), shown in Figure 4A, EGS^{CAT}, Sequence ID No. 8),
15 can direct the specific cleavage of CAT mRNA by human RNase P *in vitro* or *in vivo* cells in tissue culture. However, the cleavage reaction is inefficient compared to cleavage of natural tRNA precursor substrates.

The proposed secondary structure of a complex of CAT mRNA (Sequence ID No. 7) and EGS^{CAT} (Sequence ID No. 8) resembles
20 the tRNA cloverleaf structure, but it includes sequences not normally found in the tRNA from which it was originally derived, namely, tyrosyl tRNA (tRNA^{Tyr}) (Sequence ID No. 6) of *Escherichia coli*. To ensure that appropriate tertiary interactions that facilitate the process of enzyme-substrate recognition exist in the complex, parts of the EGS that
25 participate in tertiary interactions in the analogous tRNA structures were changed in two ways. First, four nucleotides in the equivalent of the T loop and five in the equivalent of the variable loop were randomized by incorporation of equimolar quantities of the deoxyribonucleotides dA, dG, dC and T into a DNA template to yield an initial population of 2.6×10^5
30 sequence variants. Second, during each round of selective amplification, random mutations were introduced by performing polymerase chain reaction (PCR) at an error rate of approximately 0.1 percent per

nucleotide incorporated, using the method of A. Beaudry and G.F. Joyce, *Science* 257: 635 (1992). The mRNA for the gene for chloramphenicol acetyltransferase (CAT) can be easily manipulated on plasmids and the enzymatic activity is readily expressed in tissue culture cells so it was

5 used as a target substrate. As demonstrated below, an EGS can target CAT mRNA (Sequence ID No. 7) for specific cleavage by human RNase P. Figure 4A shows a complex in which an EGS, EGS^{CAT} (Sequence ID No. 8), could base-pair with nucleotides 67 to 79 of CAT mRNA (Sequence ID No. 7), where the first nucleotide of the translation

10 initiation codon is numbered 1, and direct human RNase P to cleave that mRNA at nucleotide 67. The EGS^{CAT} (Sequence ID No. 8) construct was derived from the *E. coli* tRNA^{Tyr} (Sequence ID No. 6) gene where the first eighteen nucleotides from the 5' terminus have been deleted and the sequences on the D-loop and acceptor-loop have been changed to make

15 base-pairs with CAT mRNA. The EGS^{CAT} (Sequence ID No. 8) fused upstream with a T7 promoter was cloned into a pUC19 vector. The EGS^{CAT} RNA (Sequence ID No. 8) was prepared through *in vitro* transcription with T7 RNA polymerase.

The *Hind*III-*Bam*HI fragment of CAT gene (pCATTM, Promega) was cloned in pGem-2. The plasmid was truncated with *Eco*RI and a 260 nucleotide-long transcript was obtained by *in vitro* transcription with T7 RNA polymerase in the presence of [α -³²P]GTP. The EGS sequence was synthesized by polymerase chain reaction, using the *E. coli* tRNA^{Tyr} gene as template, with oligonucleotide GCCAAACTGAGCAG

25 ACTC (Sequence ID No. 12) and GCGCggtaccAAAAATGGTGAGG CATGAAGG (Sequence ID No. 13). The bold letters in the oligonucleotide sequences indicate the bases needed to make base pairs to CAT mRNA, the underlined letters indicate the sequence complementary to the transcription termination signal, and the lower case letters shows an

30 extra linker sequence. The sequence GCGC at the 5' end of the second oligonucleotide are extra nucleotides. The PCR fragment was digested with *Hind*III and cloned into pUC19 with a T7 promoter upstream from

the EGS sequence. The EGS^{CAT} RNA (Sequence ID No. 8) was transcribed with T7 RNA polymerase after the plasmid was linearized with *Dra*I. A mixture of unlabelled and [α -³²P]GTP labelled CAT mRNA fragment, 0.2 pmole in total, was mixed with the EGS^{CAT} RNA (Sequence ID No. 8) in amounts of 4 pmole, 1 pmole, 1 pmole, 0.4 pmole and 0.2 pmole. Each mixture was incubated at 37°C in 50 mM Tris-Cl (pH 7.5), 100 mM NH₄Cl, and 25 mM MgCl₂ with enzyme for one hour. The reaction was stopped by addition of an equal volume of dye solution with excess EDTA and then subjected to a 5% polyacrylamide-7 M urea gel. CAT mRNA alone was incubated without and with enzyme and loaded on the gel.

Primer extension analysis determining the precise site of EGS^{CAT}-directed cleavage by human RNase P was conducted as follows. A reverse transcription reaction was performed on the uncleaved and cleaved CAT mRNA using an oligodeoxyribonucleotide GGCCGTA ATATCCAGCTGAACGG (Sequence ID No. 14), complementary to nucleotides 129 to 107 of CAT mRNA. The reaction was incubated in 100 mM Tris-Cl (pH 8.3), 10 mM KCl, 6 mM MgCl₂, 10 mM DTT and 2 units AMV reverse transcriptase at 46°C for 2 hours. Labelled G, A, U, C were used as reference analyses of DNA sequences corresponding to the CAT mRNA template.

The precise site of cleavage of CAT mRNA was determined by primer extension analysis using an oligodeoxyribonucleotide primer complementary to nucleotides 129 to 107 of the RNA, showing that the cleavage occurs between nucleotides 66 and 67, as expected.

The results of the EGS^{CAT} RNA-directed cleavage of CAT mRNA were analyzed by gel electrophoresis. In the presence of EGS^{CAT} (Sequence ID No. 8) molecules, CAT mRNAs were cleaved to give rise to two products with the expected size. Analysis of the products of the reaction showed that the end groups contained 5' phosphoryl and 3' hydroxyl termini, the same as those normally generated by RNase P. The

results show conclusively that the specific cleavage of CAT mRNA is due to an EGS-directed RNase P hydrolytic reaction.

For up to five-fold molar excess of EGS^{CAT} RNA to mRNA, the cleavage efficiency is proportional to the amount of EGS^{CAT} added.

- 5 However, more than ten-fold excess of EGS^{CAT} molecules caused a decrease in the cleavage efficiency. One explanation for this is that EGS^{CAT} alone inhibits the enzymatic activity by competing with the mRNA-EGS complex for the enzyme. The reaction proceeds in a linear fashion for more than 3 hours at 37°C. Denaturation and reannealing of
10 the oligonucleotides in the target complex did not improve the efficiency of cleavage. The reaction has an absolute requirement for Mg²⁺ with an optimal concentration of 25 mM, in contrast to that with tRNA^{Tyr} precursor as substrate, which has an optimal Mg²⁺ concentration of between 2 and 10 mM.

15

Example 3: Inhibition of Expression of CAT activity in green monkey CV-1 cells by EGS^{CAT}.

- In order to test whether the EGS can function *in vivo*, the EGS^{CAT} sequence (Sequence ID No. 8) was inserted downstream of a
20 mouse U6 snRNA gene promoter in a BLUESCRIPTTM (Stratagene, La Jolla, CA) vector forming pEGS^{CAT}. The EGS^{CAT} sequence (Sequence ID No. 8) can be transcribed by RNA polymerase III and the transcription can terminate at a T₃ cluster following the EGS sequence in either S100 extract or living cells. Green monkey fibroblast cells CV-1 were
25 cotransfected with pCAT and pEGS^{CAT} plasmids. After transfection with pCAT, which encodes the CAT gene, and pEGS^{CAT}, cells were harvested and CAT activity was assayed.

- CV-1 cells were maintained in Dulbecco's modified Eagle medium that contained 10% fetal calf serum. One day prior to
30 transfection cells were split 1:10 and plated in 60 mm Petri plates. Two hours prior to transfection cells were fed with 4.5 ml of fresh medium with 10% fetal calf serum. Transfection was performed by the calcium

phosphate precipitation procedure using 2.5 μ g of pCAT DNA and various amounts of pEGS^{CAT} DNA, ranging from one to 6.25 μ g. Twenty-four hours after transfection, cells were harvested and cell extracts were assayed for CAT activity.

- 5 The extract from cells cotransfected with EGS^{CAT} construct apparently decreased the conversion of chloramphenicol to its acetylated forms. The degree of inhibition was measured quantitatively by counting of the spots excised from a TLC plate. EGS^{CAT} cotransfection produced greater than 50% inhibition compared to the control with no EGS^{CAT} cotransfection. There was substantial loss of ability to inhibit CAT expression when a higher ratio of pEGS^{CAT} to pCAT was introduced. Similar experiments yielding approximately 70% of CAT activity have also been performed with human cells in tissue culture.
- 10

15 **Example 4: Preparation of EGS with altered sequence that enhance degradation of target RNA by RNase P.**

- As explained in detail below, two classes of EGS were designed, as shown in Figures 4A, 4B, 4C, 4D, 4E, and 4F. The first class involves deletions of large segments of the EGS as suggested in
- 20 Example 1 and described in Examples 2 and 3. For example, it has been found that the anticodon loop and stem can be deleted, as shown in EGS^{CAT} Δ AC (Sequence ID No. 10) of Figure 4C and 4F, and the EGS can still promote cleavage by RNase P of a target RNA (CAT mRNA containing Sequence ID No. 7 in Figures 4C and 4F). The anticodon
- 25 loop and part of the variable loop can alternatively be deleted from the EGS, and the EGS will promote cleavage with greater efficiency than the parent EGS molecule. The most efficient EGS of this deletion class was the one in which the anticodon stem and loop was deleted. This EGS^{CAT} Δ AC (Sequence ID No. 10) promoted cleavage by human RNase
- 30 P of a target mRNA (CAT mRNA) at a rate 10-fold higher than the parent EGS. The deletion of both variable and anticodon stems and loops, however, does not yield a more efficient EGS, although the

variable loop can consist of only one or two nucleotides and still be highly efficient.

The second class of EGSs have changes in both the equivalent of the T loop, the variable loop, and the anticodon stem of the tRNA-like segment of the EGS. Three such EGSs, described below, are EGS 6,
 5 EGS 8 and EGS 9 (Sequence ID No. 9, Figures 4B and 4E). As shown below, EGS 9 is the most efficient of the EGSs in these examples and directs cleavage by human RNase P of a target RNA (CAT mRNA) at a rate approximately fifty to one hundred fold greater than the parent EGS.

10 These results apply to relative rates of cleavage at a particular site in a target mRNA. The absolute rates of cleavage at any particular site still depend on access of the EGS to that particular site.

Example 5: Preparation of RNAs with randomized nucleotides.

15 **A. Preparation of Chimeric Covalently Linked mRNA-EGS Substrate.**

The procedure to select for EGSs that are more efficient in guiding RNase P to the target CAT mRNA involves the synthesis of a population of chimeric, covalently linked mRNA-EGS substrates which
 20 are then run through cycles of *in vitro* mutation and selection for molecules which can serve as substrates for RNase P. Double-stranded DNA templates were made by annealing of and enzymatically extending two overlapping synthetic oligonucleotides: TAATACGACTCACTATA
 25 GAACATTTTGAGGCATTTTCAGTCAGTTGGCCAAACTGAGCAGAC
 (SEC-1A, Sequence ID No. 2) and TGGTGAGGCATGAAGGNN
NNGAACCTTCNNNNNGCAGATTAGAGTCTGCTCAGTTTGGCC
 (SEC-1B, Sequence ID No. 3), where the complementary sequences are underlined and the randomized nucleotides (N) were introduced during
 30 their machine synthesis by incorporating equimolar quantities of four nucleotides. These sequences create a chimeric tRNA gene which contains sequences from CAT mRNA and tRNA^{Tyr} from *E. coli* as well as nine nucleotides (N) that are randomized. A promoter for T7

bacteriophage RNA polymerase is included in SEC-1A (Sequence ID No. 2). The extension was carried out with AMV reverse transcriptase at 46°C for two hours. Variant RNA pools were prepared by transcription with T7 polymerase in 40 mM Tris.Cl, pH 7.9, 6 mM MgCl₂, 10 mM dithiothreitol 2 mM spermidine 1 mM NTPs containing 20 μCi [α -³²P]GTP at 37°C.

B. The Selection Procedure.

The general selection scheme is described above. One of the template-creating oligonucleotides (SEC-1A, Sequence ID No. 2) was also used as the 5' primer for the polymerase chain reaction (PCR) in order to allow restoration of the T7 promoter sequence and the leader sequence of the chimeric RNA for the next cycle of selection. The stringency of selection was increased at each cycle by reducing the amount of enzyme and the time allowed for the cleavage reaction, such that only those substrates that were cleaved rapidly by the enzyme were selected.

In the first three rounds of selection, RNA substrates were digested with 3.6 units of human RNase P, purified through the glycerol gradient step described by Yuan and Altman, *Proc. Natl. Acad. Sci. USA* 89: 8006-8010 (1992) and Bartkiewicz *et al.*, *Genes and Dev.* 3: 488 (1989)) in 50 mM Tris-Cl (pH 7.5), 10 mM MgCl₂, and 100 mM NH₄Cl at 37°C for 2 hours. One unit of human RNase P is defined as that amount of enzyme that cleaves 1 pmol of precursor to tRNA^{Tyr} from *E. coli* in 30 min at 37°C. For assays in subsequent rounds of selection, the amount of enzyme was reduced, and the incubation time was shortened so that less than 20 percent of the substrate was cleaved.

Cleavage products were separated from uncleaved substrates by electrophoresis on an eight percent polyacrylamide-7 M urea gel. RNA was extracted from the gels by the crush and soak method.

The purified cleavage product RNAs were reverse transcribed and amplified by PCR with SEC-1A (Sequence ID No. 2) and SEC-1C (TGGTGAGGCATGAAGG, Sequence ID No. 15) as primers with a Perkin Elmer RNA PCR kit. The double-stranded DNA generated by

PCR regained the T7 promoter sequence and the leader sequence from the sequence in the primer SEC-1A, and it was then used as a template for transcription of RNA for the next round of selection.

5 C. Characterization of the selected RNAs and EGS RNA derived from them.

After eight cycles of selection the resulting pool of double-stranded DNAs were cloned into the BLUESCRIPT™ vector (Stratagene, La Jolla, CA) vector. Eighteen plasmid DNAs were sequenced using Sequenase 2.0 (U.S. Biochemicals, Cleveland, OH).

10 In order to test the abilities of EGSs derived from the individual variants selected above for CAT mRNA targeting, sequences corresponding to the EGS segment of each chimeric tRNA were amplified by PCR using primers SEC-1A (Sequence ID No. 2) and SEC-1T (TAATACGACTCACTATAGGCCAACTGAGCAGAC, Sequence ID
15 No. 16), which contains a promoter sequence for T7 polymerase, and RNAs were transcribed with T7 RNA polymerase. EGS-directed CAT mRNA cleavage was assayed in 10 µl of 50 mM Tris-Cl, pH 7.5, 10 mM MgCl₂, 100 mM NH₄Cl containing 0.25 pmol (1000 cpm) of substrate RNA and 1 or 5 pmol of EGS RNAs. Reaction mixtures were incubated
20 at 37°C for 30 min with 10 units of RNase P from HeLa cells, followed by electrophoresis in 5% polyacrylamide/7 M urea gels.

The gels showed a species of RNA migrating in the position expected for cleavage of substrate RNA in those lanes where the newly selected EGSs have been included in the reaction mixtures.

25 D. Sequence Analysis of Randomized EGSs.

Sixteen individual clones were sequenced. The sequence of the anticodon stem/loop, the variable (V) loop, and the T stem/loop are shown in Figure 6. From the sequence that had been randomized in the T-loop, two particular sequences were most frequently selected:
30 UUCGUGC, found in seven clones, and UUCGCCC, also found in seven clones. The T loop sequences of the two remaining clones contained single transition mutations of the two major sequences, UUCGUCC and

UUCACCC. By contrast, no significant sequence-related bias was seen in the sequence of five nucleotides in the variable loop.

In addition to the sequences in the T and variable loops that were selected from the totally randomized sequence, a considerable number of mutations were introduced into the EGS in the chimeric substrates as a consequence of the conditions for PCR. Some of these mutations were beneficial and, therefore, the sequences that included them were selected and accumulated. In almost all of the individual selected clones, the integrity of base-pairing in the anticodon stem was disrupted.

Table 1 shows the partial sequences of the anticodon loop/stem region, the variable (V) loop, and the T stem/loop region of the parent chimeric mRNA-EGS^{CAT} substrate (P), nucleotides 28 through 88 of Sequence ID No. 1. Table 1 also shows partial sequences of EGS segments of some individual chimeric mRNA-EGS substrates, nucleotides 1 through 60 of Sequence ID No. 9 and Sequence ID Nos. 17 through 31, which were obtained as a result of the *in vitro* selection procedure. Nucleotides that differ from the parent chimeric mRNA-EGS^{CAT} substrate sequence are given in bold letters and are underlined. Hyphens indicate deletions. The remaining part of the sequence of the mRNA-EGS chimeric substrates is shown in Figure 4A. Numbering of the partial sequences is not uniformly consecutive, since some clones did not have appropriate inserts and only sixteen selected sequences are listed.

Table 1:

Substrate # (ID No.)	D stem/ loop	Anticodon stem/loop	V loop	T stem/ loop	AA stem
19	30	47	52	69	
SCE-1	GGCCAAACUGA	GCAGACUCUAAAUCUGC	NNNN	GAAGGUUCNNNNCCUUC	AUGCCUCACCA
(28-88 of 1)					
#1 (17)	GGCCAAACUGA	GCAGACUCUAAAUCGGC	CCUUC	GAAGGUUCGCCCCUUC	AUGCCUCACCA
#2 (18)	GGCCAAACUGA	GCAGACUCUAAAUCUGC	ACGAGA	GAAGGUUCGUGCCUUC	AUGCCUCACCA
#4 (19)	GGCCAAACUGA	GCAGACUCUAAAUCGGC	<u>CUAAC</u>	GAAGGUUCGCCCCUUC	AUGCCUCACCA
#5 (20)	GGCCAAACUGA	GCAGACUCUAAAUC-UGC	<u>CCAAC</u>	GAAGGUUCACCCUUC	AUGCCUCACCA
#6 (21)	GGCCAAACUGA	GCAGACUCUAAAUC--C	<u>ACCAA</u>	GAAGGUUCGUGCCUUC	AUGCCUCACCA
#8 (22)	GGCCAAACUGA	GCAGACUCUAAAUC-CUCC	<u>UCCCA</u>	GAAGGUUCGUGCCUUC	AUGCCUCACCA
#9	GGCCAAACUGA	GCAGACUCUAAAUC-GC	<u>AAACG</u>	GAAGGUUCGUGCCUUC	AUGCCUCACCA
(1-60 of 9)					
#10 (23)	GGCCAAACUGA	GCAGACUCUAAAUCGGC	<u>CUACG</u>	GAAGGUUCGCCCCUUC	AUGCCUCACCA
#11 (24)	GGCCAAACUGA	GCAGACGCUAAAUCUAC	<u>CCCGU</u>	GAAGGUUCGUGCCUUC	AUGCCUCACCA
#12 (25)	GGCCAAACUGA	GCAGACUCUAAAUCUUC	<u>CACCA</u>	GAAGGUUCGCCCCUUC	AUGCCUCACCA
#13 (26)	GGCCAAACUGA	GCAGACUC-AAAUCUGGC	<u>CAUUC</u>	GAAGGUUCGCCCCUUC	AUGCCUCACCA
#14 (27)	GGCCAAACUGA	GCAGACUCUAAAUC-GC	<u>AGUGU</u>	GAAGGUUCGUGCCUUC	AUGCCUCACCA
#15 (28)	GGCCAAACUGA	GCAGACUCUAAAUCAGC	<u>GCGUG</u>	GAAGGUUCGUGCCUUC	AUGCCUCACCA
#16 (29)	GGCCAAACUGA	GCAGACUCUAAAUCGGC	<u>CGCAC</u>	GAAGGUUCGCCCCUUC	AUGCCUCACCA
#17 (30)	GGCCAAACUGA	GCAGACUAAAUCUUC	<u>ACGAG</u>	GAAGGUUCGCCCCUUC	AUGCCUCACCA
#18 (31)	GGCCAAACUGA	GCAGACCCUAAAUCUGC	<u>CCCCG</u>	GAAGGUUCGUGCCUUC	AUGCCUCACCA

E. Synthesis and analysis of individual EGS RNAs based on selected chimeric substrates.

The best chimeric substrate selected was clone 9 whose corresponding RNA sequence is partially shown in sequence 9
5 (nucleotides 1 through 60 of Sequence ID No. 9) in Table 1. The clone 9 chimeric substrate was cleaved about 5.5 times more efficiently than the parent, non-randomized, chimeric substrate, mRNA-EGS^{CAT} chimera. Using the sequences of the selected chimeric substrates, nine individual corresponding EGS RNAs were synthesized: EGS-1 (Sequence ID No.
10 17), EGS-4 (Sequence ID No. 19), EGS-5 (Sequence ID No. 20), EGS-6 (Sequence ID No. 21), EGS-8 (Sequence ID No. 22), EGS-9 (nucleotides 1 through 60 of Sequence ID No. 9), EGS-12 (Sequence ID No. 25), EGS-14 (Sequence ID No. 27) and EGS-18 (Sequence ID No. 31), in order to probe the function of these selected EGSs in directing RNase P
15 to the target CAT mRNA.

Each of the individual EGS RNAs was mixed with ³²P-labeled CAT mRNA and the mixtures were then exposed to RNase P. Every selected EGS RNA increased the initial rate of the cleavage reaction, as measured during the linear phase of the reaction, over that with EGS^{CAT},
20 and cleavage occurred at the expected site in the target mRNA, as demonstrated by Figure 5.

The most dramatic improvement in rates occurred with the EGS sequence based on clone 9, EGS 9 (nucleotides 1 through 60 of Sequence ID No. 9), which directed cleavage of the CAT mRNA at an
25 overall rate more than 30 times faster than that observed with EGS^{CAT} in the complex. The three most efficient EGSs tested, derived from clone 6, clone 8 and clone 9, all had a common sequence, UUCGUGC, in the T loop.

F. Secondary structural analysis of mRNA-EGS complexes.

30 The proposed secondary structure of the complex of CAT mRNA and EGS 9 (Sequence ID No. 9) (Figure 4B) can be compared with the parent CAT mRNA-EGS^{CAT} complex (Figure 4A). The

structures of CAT mRNA-EGS complexes were confirmed in part by partial digestion with RNases T1 and T2 under conditions that allowed formation of the mRNA-EGS complex and identification of single-stranded regions in RNA. Partial digestion of CAT mRNA-EGS complexes with RNases T1 and T2 (Pharmacia) were performed in RNase P assay buffer (50 mM Tris-HCl (pH 7.5), 10 mM MgCl₂, and 100 mM NH₄Cl). The reaction mixture contained substrate RNA labeled with [³²P] at its 5' terminus (2000 cpm), 0.2 mg/ml of rat 5S RNA as carrier, and three different concentrations, 2 x 10⁻⁴, 1 x 10⁻³ and 5 x 10⁻³ units/ml, of RNase T1 or RNase T2. Reactions were incubated at room temperature for 5 min. The samples were analyzed by PAGE in 12.5% sequencing gels. Double-stranded regions were identified by digestion with cobra venom nuclease (Pharmacia, Alameda, CA). Conditions for digestion with cobra venom nuclease were as described above for RNase T1 and T2 except that incubation was at 37°C.

The sites of cleavage by the various nucleases in CAT mRNA-EGS^{CAT} complex are indicated by solid arrows in Figures 4A and 4D. The tRNA domain in this structure is very similar to that found in a natural tRNA. The results obtained with cobra venom nuclease indicate that the first few nucleotides in the analog of the D loop are involved in a tertiary interaction, presumably with nucleotides in the variable loop (Figure 4D). This interaction is either absent or much less extensive in the complex with EGS 9 (Sequence ID No. 9) (Figure 4E) and, indeed, the same region in the analog of the D loop is susceptible to attack by RNases T1 and T2, an indication that it is in a single-stranded conformation (Figure 4B). This result, together with the appearance of new sites of susceptibility to attack by cobra venom in the anticodon loop of EGS 9 confirm that this EGS endows the complex of the EGS and CAT mRNA with new tertiary interactions that enhance the rate of cleavage by RNase P. The results of cleavages by nucleases also confirmed that the anticodon stem in EGS 9 was disrupted as a result of a single nucleotide deletion. EGSs 9 and 14 are not similar to each other in

terms of their efficiency in directing RNase P to a target substrate. The only difference in their nucleotide sequences is in the variable loop, as shown by Table 1. This difference alone must account for the relative inefficiency of EGS 14 in terms of targeting ability. Furthermore, digestion with RNase T1 of the complex that contained EGS 9 revealed strong protection of the last nucleotide, G, in the variable loop from attack by RNase T1, as shown by Figure 4B. The role of this G may be similar to that played by nucleotide 57 in tertiary interactions in tRNA molecules; namely, the G may form hydrogen bonds with a nucleotide in the CAT mRNA sequence to ensure folding of the "tRNA" domain of the mRNA sequence.

G. Anticodon stem and loop of an EGS decrease cleavage rate of RNase P and are dispensable.

An independent study of recognition of precursor tRNA substrates by human RNase P showed that the anticodon stem and loop form a dispensable structural feature in the recognition of substrates by human RNase P. To determine whether the anticodon stem and loop may act in a negative fashion on the overall rate of cleavage of target RNA by RNase P, two more EGS RNAs were constructed, one being a deletion mutant that lacked the equivalent of the anticodon stem and loop, EGS^{CAT}ΔAC (Sequence ID No. 10), as compared to the parent EGS^{CAT} (Figure 4A), and the other, EGS 19, being a derivative of EGS 9 in which the structure of the anticodon stem was restored.

DNA coding for EGS^{CAT}ΔAC was synthesized by PCR with pEGS^{CAT} DNA, described by Yuan *et al.*, *Proc. Nat. Acad. Sci. USA* 89: 8006 (1992), as template. Oligonucleotide EC-1ΔAC (GCCAAACTG ACGTCATCGACTTCG, Sequence ID No. 32) and M13 reverse primer (AACAGCTATGACCATG, Sequence ID No. 33) were used as primers. The DNA generated by PCR was digested with *Hind*III and then inserted into pUC19 downstream from a T7 RNA polymerase promoter sequence. EGS^{CAT}ΔAC RNA (Sequence ID No. 10) was prepared by transcription *in vitro* after the new plasmid DNA had been linearized with *Dra*I. DNA

that coded for EGS 19 was synthesized by PCR procedure in a manner similar to that used for the synthesis of DNA for EGS 9, with oligonucleotides SEC-1C (Sequence ID No. 15) and SEC-1I (GTAATACGACTCACTATAGGCCAAACTGAGCAGACTCTAAATCT
5 GCAAACGGAAGGTTC, Sequence ID No. 34): the newly inserted T residue in SEC-1I is underlined. The DNA was transcribed *in vitro* with T7 RNA polymerase to give EGS 19 RNA. EGS 19 (Sequence ID No. 39) RNA differs from EGS 9 RNA only in the additional U that restores the structure of the anticodon stem.

10 In EGS^{CAT}ΔAC (Sequence ID No. 10), the length of EGS^{CAT} was reduced by 25%: the shorter deletion mutant directed cleavage of target RNA about six times more efficiently than the parent EGS, as shown by Figure 5. Restoration of the anticodon-stem structure, as in EGS 19, reduced the rate of cleavage of the target RNA with EGS 19 to
15 four times lower than that with EGS 9. These results, together with the measurement of the rates of reaction with EGSs selected *in vitro*, indicate a significant inverse correlation between the efficiency of an EGS in the cleavage reaction and the existence of an anticodon stem in EGS RNA.

H. Stability of EGS-mRNA complexes.

20 The stability of EGS-mRNA complexes was measured based on both the binding constants between the mRNA and each EGS and the dependence on Mg²⁺ ions of the cleavage reaction, presuming that relatively high concentrations of Mg²⁺ ions were needed to stabilize relatively unstable complexes. The dissociation constants (K_d) of mRNA-
25 EGS complexes were measured directly by a gel mobility shift assay in polyacrylamide gels that contained 10 mM Mg²⁺ ions according to the method of Pyle *et al.*, *Proc. Natl. Acad. Sci. USA* 87: 8187-8191 (1990). A fragment of CAT mRNA, 160 nucleotides in length, was prepared by transcription with T7 RNA polymerase in the presence of [α -³²P]GTP. 10
30 μ l EGS RNA in 2X binding buffer was heated at 80°C for 4 min before it was mixed with an equal volume of 2 nM CAT mRNA fragment in water. 1X binding buffer contains 50 mM Tris-Cl (pH 7.5), 10 mM

MgCl₂, 100 mM NH₄Cl, 3% glycerol, 0.05% xylene cyanol. The mixtures were incubated at 37°C for 20 min and immediately separated on 5% polyacrylamide gels at 9 watts. The electrophoresis buffer consisted of 36 mM Tris base, 64 mM HEPES, 0.1 mM EDTA, 10 mM MgCl₂ (pH 7.5 without any adjustment). Quantitation of free target RNA and of the complex was performed with a Betascope (Betagen, Waltham, MA). The free energies of binding were determined from the equation $\Delta G^\circ = -RT \ln(1/K_d)$, where R = 0.00198 kcal/mol and T = 310.15°K.

The dissociation constants for selected EGSs are shown in

Table 2.

Table 2: Dissociation constants of EGSs.

Substrates	K _d (nM)	K _m (nM)	V _{max} (nmol/min) x10 ⁻⁵	V _{max} /K _m x10 ⁻⁶
pTyr		10	2.9	2.90
CAT mRNA				
+ EGS ^{CAT}	880	120	2.9	0.24
(Sequence ID No. 8)				
+ EGS ^{CAT} ΔAC	20	150	11.4	0.76
(Sequence ID No. 10)				
+ EGS-5	210	125	16.3	1.30
(Sequence ID No. 20)				
+ EGS-8	25	125	21.3	1.70
(Sequence ID No. 22)				
+ EGS-9	78	125	30.0	2.40
(nucleotides 1-60 of Sequence ID No. 9)				
+ EGS-19 710	130	12.5	0.96	

Table 2 shows the kinetic parameters of EGS-directed cleavage of CAT mRNA in vitro by RNase P from HeLa cells. K_d refers to measurements of the dissociation constant for binding of EGS to CAT mRNA. The other parameters were determined in standard assays of enzyme kinetics. V_{max} is the value obtained with 0.5 ml (0.6 units) of human RNase P. pTyr refers to the precursor to tRNA^{Tyr} from *E. coli*.

The results shown in Table 2 indicate that the K_d values of *in vitro* selected EGSs are 4 to 40 times lower than that of the parent EGS.

Thus, the selected EGSs had higher affinity for the target RNA than did EGS^{CAT}. The chimeric substrate derived from clone 9 was cleaved by RNase P at a rate only about 1.5 times faster than was the target in the mRNA-EGS 9 complex, an indication that the ability of the EGS to bind
5 tightly to the target RNA in solution must be a critical determinant in the efficiency of the substrate complexes.

The differences in K_d values between complexes with EGS^{CAT} (Sequence ID No. 8) and the selected EGSs correspond to the contribution of -1 to -2.4 kcal/mol to the free energy of binding (ΔG°) with selected
10 EGSs (ΔG°) is -8.5 kcal/mol for the complex with EGS^{CAT}, -10.1 for the complex with EGS 9 (Sequence ID No. 9) and -10.9 for the complex with EGS^{CAT} Δ AC (Sequence ID No. 10), thus revealing new interactions in the selected EGS-mRNA complexes.

Deletion of the anticodon stem from EGS^{CAT} resulted in the K_d
15 for EGS^{CAT} Δ AC being 44 times lower, and restoration of the stem of EGS 9 (EGS 19) resulted in a K_d that was ten times higher than that for EGS 9 and was close to that of EGS^{CAT} (Table 2). Thus, the intact anticodon stem stabilized a conformation of the EGS that could bind as strongly to the target RNA as does an EGS with no organized anticodon stem.
20 Accordingly, the enhancement of the ability of the selected EGSs for targeting RNA can be assigned, in part, to the increase in the strength of their binding to target RNA.

Cleavage of mRNA directed by the original EGS^{CAT} requires Mg^{2+} ions with an optimum concentration of 25 mM. By contrast, the
25 reaction with certain selected EGSs, namely, EGS 6 (Sequence ID No. 21) and EGS 9 (nucleotides 1 through 60 of Sequence ID No. 9), proceeds optimally in 2 to 10 mM Mg^{2+} ions. This latter concentration is close to the optimal concentration of Mg^{2+} ions for processing of tRNA precursors by human RNase P, reported by Doersen *et al.*, *J. Biol.*
30 *Chem.* 260: 5942 (1985). Since high concentrations of Mg^{2+} ions are especially effective in the neutralization of repulsion between adjacent regions of the phosphate backbone and the stabilization of RNA folding,

the results indicate that the selected EGSs can achieve the appropriate folded structures in the complex with target RNAs without the aid of high concentrations of Mg^{2+} ions.

A kinetic analysis was performed to determine the Michaelis
5 constant (K_m) and the maximum velocity (V_{max}) of the enzymatic
reactions. The cleavage of the precursor to tRNA^{Tyr} from *E. coli* and of
CAT mRNA in mRNA-EGS complexes was assayed at various substrate
concentrations both above and below the K_m for these substrates.
Aliquots were withdrawn from reaction mixtures at regular intervals and
10 analyzed on polyacrylamide-urea gels. Values of K_m and V_{max} were
obtained from Lineweaver-Burk double-reciprocal plots. The effective
concentrations of substrate used were calculated as the concentration of
the complex of target mRNA with EGS as determined from the K_d values
shown in Table 2. The K_m for the precursor to tRNA^{Tyr} (pTyr) is 10 nM
15 with human RNase P, whereas the K_m value of the complex of mRNA-
EGS^{CAT} is twelve-fold higher (Table 2). The K_m for all the selected EGSs
tested was the same as that of EGS^{CAT}. However, the maximum velocities
of the reactions with selected EGSs were up to ten times higher than that
with the original EGS. Thus, the value of V_{max}/K_m for selected EGSs was
20 increased. The value of V_{max}/K_m of EGS 9, was for example, ten-fold
higher than that of EGS^{CAT} and was very close to that of the tRNA^{Tyr}
precursor. When the anticodon stem of EGS 9 was restored, however, as
it was in EGS 19, V_{max} fell about 2.5-fold, suggesting that the rate of
release of the product was specifically reduced by physical interactions of
25 the enzyme with an intact anticodon stem. These data show that the
enhanced abilities for targeting of the selected EGSs, as measured in the
overall rate of the cleavage reaction, were due to both enhanced affinity
of binding to substrate RNAs and to increases in the velocity of the
enzymatic reaction.

Example 6: Inhibition of viral mRNA expression with human ribonuclease P.

Herpes simplex virus was used to demonstrate that EGS can be used to target a viral gene *in vivo* to inhibit viral replication. Herpes simplex viruses are DNA-containing viruses that infect cells, induce synthesis of messenger RNAs, which are transcribed to produce enzymes related to DNA synthesis and breakdown: including thymidine kinase, DNA polymerase and a DNA exonuclease, and viral DNA and viral structural proteins are made and assembled into infectious viral particles. The structure and organization of the herpes simplex virus genome is known, for example, as reported by Roizman, *Cell* 16: 481-494 (1979). The nucleotide sequence of the thymidine kinase gene of herpes simplex type 1 was described by Wagner *et al.*, *Proc. Natl. Acad. Sci. USA* 78: 1441-1445 (1981).

An EGS, shown in Figure 6A, was designed to target the mRNA encoding thymidine kinase (TK). The target site for RNase P cleavage is about 25 nucleotides downstream from the TK translation initiation site. An EGS forming an approximately three-fourths-like tRNA was designed and shown *in vitro* to cleave the TK sequence at the proposed cleavage site. Two other EGSs, which contain a single point mutation in the T-loop (C to G) or deletion of the anti-codon region, shown in Figures 6B and 6C, respectively, were constructed based on the results observed in Example 5.

Cell lines and EGS expression vectors were then constructed. Five cell lines were constructed by transfecting plasmid DNAs into human 143TK-cells, which can be obtained from the American Type Culture Collection, Rockville, MD.

Plasmid pFL116 was constructed using pGEM-7Z (Promega, Wisconsin) and incorporating a gene for neomycin resistance (Neo), which is commercially available from Clontech or Stratagene, CA. EGS DNA (Figures 6A, 6B, and 6C) was digested with *KpnI* and inserted into the plasmid pmU6(-315/1), described by Yuan *et al.*, *Proc. Natl. Acad.*

Sci. USA 89: 8006-8010 (1992) and Das *et al.*, *EMBO* 7: 503-512 (1988), at the *Pst*I (blunted)/*Kpn*I site. This plasmid contains the promoter for the gene for U6 small nuclear RNA, a very strong promoter, and a signal for termination of transcription (T cluster) by RNA polymerase II. EGS plasmids were designated pFL104, pFL109, and pFL112, respectively. The plasmid based on sequence 9 in Table 1 was used as a control.

The cells were stably transfected using a calcium phosphate precipitation method described by Wigler *et al.*, *Proc. Natl. Acad. Sci. USA* 76: 1373-1376 (1979), with the pFL116 to yield CL116, plasmids EGS 9 and pFL116 to yield CLCAT, plasmids pFL104 and pFL116 to yield CL104, plasmids pFL109 and pFL116 to yield CL109, and plasmids pFL112 and pFL116 to yield CL112, followed by neomycin selection. Cells were cloned, expanded and RNA isolated. Both total and cytoplasmic RNA was isolated and the RNase protection method described in "Molecular Cloning: A Laboratory Manual, Second Edition" Sambrook *et al.*, (Cold Spring Harbor Laboratory Press, 1989) at pages 7.71-7.78. This is an extremely sensitive assay where digestion of RNA:RNA hybrids formed using radiolabelled probe is used to assess which cell clones express the EGS.

The results indicated that the EGSs are expressed in both nuclei and cytoplasm.

Cells were then infected with herpes simplex virus using a multiplicity of infection (MOI) of 1 to 1.5, specifically, 1 to 1.5 million viral particles/1 million cells, in order to resemble a natural infection with virus. RNA was harvested at 4, 8 and 12 hours post-infection. The internal control probe was used to detect the mRNA levels of HSV α 47 and late genes U₁₀ and U₁₁. The probe is selected to assure the detection of a high level of viral mRNA expression over the entire cycle of viral infection.

The results are shown in Figure 7. TK mRNA expressed was decreased 0% in the control CL-CAT, 30% in CL-109, 45% in CL-104, and 65% in CL-112.

Example 7: RNase P Internal Guide Sequences.**A. Construction of plasmids, catalytic RNAs and RNA substrates for studies *in vitro*.**

DNA templates for transcription *in vitro* of RNA substrates

- 5 tk7, tk46, and cat7 were constructed by annealing the T7 promoter containing oligonucleotide, OliT7 (5'-TAATACGACTCACTATAG-3') (Sequence ID No. 41) with oligo-nucleotides Olik7 (5'-CGCAGAC GGTCTATAGTGAGTCGTATTA-3') (Sequence ID No. 42), OliTK46 (5'-ACCGCCGCAGCCTGGTCGAACGCAGACGCGTGTTGATGGCAG
- 10 GGGTCTATAGTGAGTCGTATTA-3') (Sequence ID No. 43), and Olicat7 (5'-ATGCCTCGGTCCTATAGTGAGTCGTATTA-3') (Sequence ID No. 44), respectively. Plasmids pTK117 and pTK146 are derivatives of pUC19 and were described by Guerrier-Takada and Altman, *Proc. Natl. Acad. Sci. USA* **89**: 1266-1270 (1992). The DNA sequences coding
- 15 for M1 RNA (pTK117) and mutant M1 RNA with a deletion from nucleotides 167 to 377 (pTK146), are under the control of the T7 RNA polymerase promoter. The DNA templates for M1TK19, M1TK16, M1TK13, M1TK10, M1TK5, and M1CAT13 were constructed by the polymerase chain reaction (PCR) with the gene for M1 RNA as found in
- 20 plasmid pTK117 with OliT7 as the 5' primer oligonucleotide and 3' primers that contained the appropriate guide sequences. The 3' primers were OliTK19 (5'-GTGGTGTCTGCGTTCGACCAGGCTATGAC CATG-3') (Sequence ID No. 45), OliTK16 (5'-GTGGTGTCTGCGTTC GACCAGTATGACCATG-3') (Sequence ID No. 46), OliTK13
- 25 (5'-GTGGTGTCTGCGTTCTATGACCATG-3') (Sequence ID No. 47), OliTK10 (5'-GTGGTGTCTGCGTTCTATGACCATG-3') (Sequence ID No. 48), OliTK5 (5'-GTGGTGTCTGTATGACCATG-3') (Sequence ID No. 49), and OliCAT13 (5'-GTGGTGAGGCATTTCAGTTATGAC CATG-3') (Sequence ID No. 50). The 3' proximal sequences of 10 no
- 30 serve as the primers for the PCR with the pUC19 sequence. The underlined sequences and the bold sequences correspond to the 3'CCAC sequence and the guide sequences, respectively. The DNA template for

Δ M1(167-377)TK13 RNA was constructed by PCR with the sequence for M1 RNA in pTK146 and primers OliT7 and OliTK13. The DNA template for EGS TK16 was constructed by PCR with the 5' primer OliT7 and the 3' primer OliTK16 for DNA template pUCT7, which was
5 derived from pUC19 by insertion of a T7 promoter sequence into the *Bam*HI site. The CAT mRNA fragment of 550 nt was synthesized with T7 RNA polymerase and *Eco*RI-digested pCAT-1 plasmid DNA (Promega Inc.) while the TK mRNA fragment of 450 nt was synthesized from plasmid pTK101 DNA in which the sequence for TK mRNA is under
10 control of the T7 promoter.

B. Assays for cleavage by M1GS RNA.

RNA enzyme (20 nM) and substrate (50 nM), either uniformly labeled with [α - 32 P]GTP or 5' end-labeled with [γ - 32 P]ATP, were incubated for 30 mins at 37°C or 50°C in buffer A (50 mM Tris, pH 7.5,
15 100 mM NH₄Cl, 100 mM MgCl₂) or buffer B (50 mM Tris, pH 7.5, 100 mM NH₄Cl) that contained MgCl₂ at various concentrations. Reactions were stopped by the addition of 8 M urea and the cleavage products were then separated on either 15% or 20% polyacrylamide gels that contained 8 M urea. C5 protein and human RNase P protein were purified from *E.*
20 *coli* and HeLa cells, respectively, as described previously by Vioque *et al.*, *J. Mol. Biol.* 202: 835-848 (1988), and Bartkiewicz *et al.*, *Genes & Dev.* 3: 488-499 (1989). The RNase P holoenzyme from *E. coli* was assembled by mixing M1 RNA and C5 protein at a molar ratio of 1:20.

Assays to determine kinetic parameters under single- and
25 multiple-turnover conditions were performed as described previously by Guerrier-Takada *et al.*, *Cell* 35: 849-857 (1983), and Liu and Altman, *Cell* 77: 1083-1100 (1994). Cleavage was assayed at various concentrations of substrate, in 2- to 20-fold excess over enzyme concentration, both above and below the K_m for the substrate. Aliquots
30 were withdrawn from the reaction mixtures at regular intervals and the cleavage products were separated in polyacrylamide-urea gels. quantitation was carried out with a phosphorimager (Molecular

Dynamics). The values of K_m and k_{cat} were obtained from Lineweaver-Burk double reciprocal plots. In single turnover experiments, trace amounts of substrates were used and the concentrations, 1 nM, were much lower than the K_m (> 80 nM). The concentration of enzyme ranged from 5 nM to 200 nM. The observed rate of cleavage (k_{obs}) was determined, and the value of k_{cat}/K_m was obtained from the equation $k_{cat}/K_m = k_{obs}/[E]$, where $[E]$ is the concentration of the enzyme.

C. Viruses, cells and antibodies.

The properties of HSV-1(F), a prototype of human herpes simplex virus 1 have been described by Ejercito *et al.*, *J. Gen. Virol.* **2**: 357-364 (1968). The retroviral vector LXSN, retroviral packaging cell lines, PA317 (amphotropic) and ψ CRE (ectopic), and their maintenance and propagation were described by Miller and Rosman, *BioTechniques* **7**: 980-990 (1989), and Danos and Mulligan, *Proc. Natl. Acad. Sci. USA* **85**: 6460-6464 (1988). The rabbit polyclonal antibody against thymidine kinase of HSV-1 (F) was described by Liu and Summers, *Virology* **163**: 638-642 (1988), and the mouse monoclonal antibody MCA406 against HSV-1 ICP35 protein was described by Liu and Roizman, *J. Virol.* **65**: 206-212 (1991), and purchased from Harlan Bioproducts for Sciences Inc. (Indianapolis, IN). The secondary antibodies used in the Western blots, anti-rabbit or anti-mouse IgGs conjugated with horseradish peroxidase, were purchased from Vector Laboratories Inc. and Bio-Rad Inc., respectively.

D. Construction of plasmids for studies *in vivo*.

The promoter sequence for U6 small nuclear RNA and a signal for the termination of transcription (T cluster) by RNA polymerase III, described by Yuan *et al.* (1992), were inserted into the *EcoRI* site for the retroviral vector LXSN, described by Miller and Rosman (1989), to create pRV0. Retroviral constructs PB2, M1TK and Δ M1TK were constructed by placing the DNA sequence that coded for the EGS with the ability to target the mRNA for influenza viral protein PB2, m1TK13 RNA and Δ M1(167-377)TK13 RNA under the control of the U6 promoter in

plasmid pRV0, respectively. Plasmid pTK129 and pTK141 were constructed, respectively, by placing the *Bgl*III-*Mlu*I fragment (87 nt) of the HSV-1(F) *Bam*HI Q fragment and *Bam*HI-*Acc*I fragment (181 nt) of the HSV-1 (F) *Bam*HI Z fragment under the control of a phage T3 RNA polymerase promoter. These two fragments correspond to the sequences that encode the 5' sequence of TK mRNA and of the overlapping transcripts for the HSV-1 α 47, Us10, and Us11 genes, respectively, as described by McMeoch *et al.*, *J. Gen. Virol.* 69: 1531-1574 (1988).

E. Construction of cell lines.

The protocols were modified from Miller and Rosman (1989). In brief, cells were transfected with retroviral vector DNAs with the aid of a mammalian transfection kit purchased from Stratagene Inc. (La Jolla, CA). Forty-eight hours post transfection, neomycin (Gibco-BRL) was added to the culture medium at a final concentration of 600 μ g/ml. Cells were subsequently selected in the presence of neomycin for two weeks and neomycin resistant cells were cloned and allowed to proliferate in neomycin-containing medium, as described by Sambrook *et al.* (1989). The newly constructed cell lines, NB2, Δ M1TK, and M1TK, and a control cell line in which cells were transfected with LXS vector DNA, were indistinguishable in terms of cell growth and viability for up to two months. Finally, aliquots of these cells were either frozen for long-term storage in liquid nitrogen or used immediately for further studies *in vivo*.

F. Viral infection and preparation of RNA and protein extracts.

Approximately 10^6 cells in a T25 flask were either mock-infected or infected with HSV-1 in 1.5 ml of Medium 199 (M199; GIBCO) supplemented with 1% fetal calf serum. After 2 hours of exposure of cells to virus at 37°C, the medium was replaced with Dulbecco's modified Eagle's medium supplemented with 5% fetal bovine serum. The cells were incubated for 11 hours before harvesting for isolation of viral mRNA and/or protein. RNA and protein extracts were prepared from cells that had either been mock-infected with HSV-1, as

described previously by Jenkins and Howett, *J. Virol.* 52: 99-107 (1984), and Liu and Roizman (1991).

G. RNase protection assay for viral mRNA expression.

The RNA probes used to detect TK mRNA and the transcripts
5 of the α 47, Us10, and Us11 genes were synthesized *in vitro* with T3
RNA polymerase (Promega, Madison, WI) from the DNA templates
pTK129 and pTK141, respectively, that had been linearized with *Eag*I.
RNase protection assays were performed as described previously by Yuan
et al. (1992). The protected RNA products were separated in 8 M urea
10 8% polyacrylamide gels and quantitated with a phosphorimager.
quantitation was performed in the linear range of RNA detection.

H. Electrophoretic separation and staining with antibodies of polypeptides from infected cells.

Denatured, solubilized polypeptides from cell lysates were
15 separated on SDS-9% [v/v] polyacrylamide gels. The separated
polypeptides were transferred electrically to nitrocellulose membranes and
allowed to react in an enzyme-linked immunoassay with antibodies against
either mouse or rabbit IgG that had been conjugated with horseradish
peroxidase after reaction with antibodies against HSV-1 TK or ICP35.
20 The membranes were subsequently stained with the color-development
substrate from peroxidase substrate kit purchased from Vector
Laboratories Inc. or reacted with the chemiluminescent substrate in a
LumiGLO™ chemiluminescence kit (Kirkegaard and Perry Laboratories
Inc.) and subsequently subjected to exposure to X-ray film. Finally, the
25 amounts of TK and ICP35 protein on the membrane were quantitated by
scanning the films with a densitometer (Bio-Rad, Inc.). Quantitation was
performed in the linear range of protein detection.

I. Cleavage of model substrates by M1GS RNA *in vitro*.

The HSV-1 TK gene has been well characterized and is
30 reviewed by Roizman and Sears, "Virology", 2nd edition, edited by
Fields *et al.*, pages 1795-1841 (Raven Press, New York, 1990).
Although the TK gene product is not essential for viral replication in

tissue culture cells, because it is so well studied we have used it as a model target for gene inactivation by M1GS RNA. DNA encoding a guide sequence (TK 13) that contains a sequence of 13 nt complementary to the 5' terminal sequence of mRNA for HSV-1 TK protein was covalently linked to the 3' end of DNA that encoded M1 RNA (Figure 9). The RNA transcript of this construct, M1TK13, cleaved target RNA, tk7, that contains 7 nt of the 5' sequence of TK mRNA and an unrelated sequence of 5 nt that serves as a leader sequence (Sequence ID No. 52). Cleavage in the target occurs at position 5, yielding two cleavage products of 5 nt and 7 nt in length, respectively. M1TK13 remains unchanged during the reaction, as expected of a true enzyme, since only labeled, full-length M1TK13 RNA is detected after incubation of the reaction mixture. Attachment of guide sequences to the 3' end, rather than to the 5' end, of M1 RNA in the construction of M1GS RNA is preferred because the presence of a 3' terminal CCA sequence in the guide sequence is important for maximum efficiency of cleavage. Furthermore, additional sequences downstream from the CCA sequence lower the rate of cleavage of substrates.

M1GS RNA only cleaves substrates that are complementary to the guide sequence. M1TK13 RNA can cleave substrate tk7 but not substrate cat7, which contains a sequence of 7 nt from the mRNA for chloramphenicol acetyltransferase (CAT). However, M1CAT13 RNA, in which the guide sequence contains a sequence of 13 nt complementary to CAT mRNA, can efficiently cleave cat7 but not tk7. Therefore, M1GS RNA appears to act as a sequence-specific endonuclease, recognizing its substrates through specific base-pairing between the GS and the target sequence, as has also been shown by Frank *et al.*, *Biochemistry* 33: 10800-10808 (1994), with different constructs. To determine whether longer RNAs can be specifically cleaved by these new RNA enzymes, uniformly [³²P]-labeled fragments of the TK (450 nts) and CAT (550 nts) mRNA sequences (Figure 10) were incubated with either M1TK13 RNA

or M1CAT13 RNA. Sequence-specific and efficient cleavage of these RNA substrates by the appropriate RNA enzymes resulted.

A set of RNA enzymes, designated M1TK19, M1TK16, M1TK13, M1TK10 and M1TK5, were constructed in which the GSs contained sequences of 19, 16, 13, 10, 5 nucleotides complementary to the TK mRNA sequence, respectively (Figure 11), in order to relate the length of the target helix to efficiency of cleavage. The substrate tk46, which contains a TK mRNA sequence of 46 nt, was cleaved by all the constructs and the cleavage site was determined to be at position G21, as expected (Figure 11). Further kinetic analysis of cleavage of substrate tk46 by M1TK16, M1TK13, M1TK10, and M1TK5 RNAs was also performed and the results are shown in Table 3.

Table 3: Kinetic parameters of reactions catalyzed by various M1GS RNA constructs

Enzyme	$K_2(\mu\text{M})$	$k_{\text{cat}} (\text{min}^{-1})$	k_{cat}/K_m	k_{cat}/K_m^*
M1TK16	0.08	0.14	1.6	2.4
M1TK13	0.09	0.21	2.3	2.4
M1TK10	0.15	0.10	0.7	0.7
M1TK5	0.22	0.03	0.1	0.1

Values of k_{cat} increased in the rank order from M1TK5 to M1TK16 RNA. Values of K_m decreased progressively with increasing length of the duplex. The overall K_m of the reaction undoubtedly involves both the substrate binding to the guide sequence (helix formation) and docking of the double helical segment in the active site of M1 RNA, a complex process. There was one exception to the rule of increasing k_{cat} with length of guide sequence: k_{cat} decreased from M1TK13 to M1TK16. This last result can be explained if release of product becomes the rate-limiting step in the reaction governed by M1TK16 RNA under single- and

multiple-turnover conditions as described by Fersht, "Enzyme structure and mechanism", 2nd edition (W.H. Freeman and Co., New York, 1985). Further analysis of the reaction with M1TK13 showed that about 5 pmol of tk46 were cleaved by 1 pmol of M1TK13 in 30 minutes, indicating that
5 the RNA enzyme turns over 5 times during the incubation period. This result is consistent with the value of k_{cat} measured by classical Michaelis-Menten kinetic analysis (see Table 3).

M1GS RNAs act more efficiently than does M1 RNA in the classic cleavage reaction in *trans*. M1TK16 cleaves tk46 at least 10 times
10 faster than does M1 RNA when using a separate guide sequence, TK16, under equivalent experimental conditions. Moreover, M1GS RNA cleaves its substrates more efficiently in buffers that more closely mimic physiological conditions in terms of magnesium ion concentration (10 mM $MgCl_2$) than does M1 RNA with a separated guide sequence in buffers
15 that contain 100 mM $MgCl_2$, an indication that the high concentrations of Mg^{2+} ions are needed to mediate binding of the substrate to M1 RNA alone, as indicated by Kazakov and Altman, *Proc. Natl. Acad. Sci. USA* 88: 9193-9197 (1991), and Smith and Pace, *Biochemistry* 32: 5273-5281 (1993). Further analysis of the cleavage reaction of M1GS RNA
20 indicated that the cleavage proceeds optimally at a temperature of 62°C, at concentrations of monovalent cations of 100 mM, and at concentrations of Mg^{2+} ions of 60-100 mM.

To prove that the catalytic activity of M1GS RNA resides in the sequence that encodes M1 RNA, a set of M1TK RNAs was
25 constructed in which each RNA had a deletion in a different region of the M1 RNA sequence but in which each had the same guide sequence, TK13. The various M1 RNA deletion mutants, for example, $\Delta 167-377$, $\Delta 1-163$, and $\Delta 65$, lack the catalytic activity needed to process pre-tRNAs, as shown by Guerrier-Takada *et al.*, *Science* 286: 1578-1584 (1989) and
30 Guerrier-Takada and Altman (1992). M1GS RNA constructs with these deletion mutants and the linked TK13 sequence were unable to cleave tk46.

Stimulation of the activity of M1GS RNA by proteins.

C5 protein, the protein subunit of RNase P from *E. coli*, increases the rate of cleavage of natural substrates by M1 RNA as described by Guerrier-Takada *et al.* (1983) and Reich *et al.*, *Science* **239**: 178-181 (1988). C5 protein also stimulates the cleavage by M1GS RNA by a factor of 30 or more. Furthermore, cleavage by M1GS RNA can be stimulated at least five-fold by a partially purified preparation of human RNase P. This enhancement in rate was anticipated as it had been previously shown that protein from a crude preparation of human (HeLa cells) RNase P can enhance the cleavage of ptRNA by M1 RNA as described by Gold and Altman, *Cell* **44**: 243-249 (1986). The rate stimulation cannot be ascribed to residual human RNase P activity as it, alone, cannot cleave substrate tk46 in complexes with catalytically inactive Δ M1(167-377)TK13 RNA. The last result is consistent with observations that a simple stem-loop structure can serve as a substrate for M1 RNA but not for eukaryotic RNase P. This was shown specifically with human and *X. laevis* RNase P by Yuan and Altman, *Science* **263**: 1269-1273 (1994), and Carrara *et al.*, *Proc. Natl. Acad. Sci. USA*. No stimulation of cleavage of substrate tk46 by M1TK13 RNA was observed when fractions devoid of human RNase P activity were used. Accordingly, it is expected that when M1GS RNA constructs are present in mammalian cells, their activity should be enhanced by endogenous proteins.

J. Expression *in vivo* of M1GS RNA.

To express M1GS RNA in mammalian cells, retroviral constructs, M1TK and Δ M1TK, were generated by cloning genes for M1TK13 and Δ M1(163-377)TK13 into the retroviral vector LXS_N under the control of a mouse U6 snRNA promoter, described by Das *et al.*, *EMBO J.* **7**: 503-512 (1988), and Yuan *et al.* (1992) (Figure 12). An additional retroviral DNA construct, NB2, containing the U6 promoter and an external guide sequence (NB2) that targets the mRNA that encodes the PB2 protein of human influenza virus, was used as a control.

The targeted cleavage site of TK mRNA expressed in cells infected with HSV-1 appears to be modifiable by dimethyl sulfate (DMS) *in vivo*, based on Peattie and Gilbert, *Proc. Natl. Acad. Sci. USA* 77: 4679-4682 (1980), Inoue and Cech, *Proc. Natl. Acad. Sci. USA* 82: 648-652 (1985), Climie and Friesen, *J. Biol. Chem.* 263: 15166-15175 (1988), and Ares and Igel, *Genes & Dev.* 4: 2132-2145 (1990). This is an indication that this site might be accessible for binding to M1GS *in vivo*. Amphotropic packaging cells (PA317) were transfected with retroviral vector DNAs to produce retroviruses that encoded the genes for M1GS RNA. Subsequently, esotropic packaging cells (ψ CRE) were infected with these retroviruses, and cells expressing the retroviruses and ribozymes were cloned. Stable expression of M1GS RNAs was demonstrated by an RNA protection assay with a probe that was complementary to the sequence of M1 RNA and RNA isolated from these cell lines. Furthermore, RNA extracted from M1TK-expressing cells can cleave substrate tk46 *in vitro* while RNA extracted from the parental ψ CRE cells, cells that expressed NB2 RNA and Δ M1TK RNA do not show cleavage activity. These results demonstrate that the M1TK RNA that is expressed in cultured cells was intact and catalytically active.

K. Inhibition of expression of TK of HSV-1 in cells that express M1GS RNA.

Cells were infected with herpes simplex virus 1 (HSV-1) at a multiplicity of infection (MOI) of 0.05 to 0.1. Levels of TK mRNA in the infected cells were determined by an RNase protection assay with an RNA probe (TK probe) that contained a sequence of 87 nt complementary to the 5' proximal sequence of TK mRNA. An RNA probe (α 47), containing a sequence of 181 nt complementary to the overlapping regions of α 47, Us10, and Us11 mRNAs encoded by HSV-1, was used to determine the levels of these latter mRNAs. The levels of the latter RNAs were used as internal controls for quantitation of expression of TK mRNA. Figure 13 graphically summarizes the results of the RNase protection experiments with both the TK and α 47 probes. A reduction of

about $80 \pm 5\%$, averaged over four experiments, in the level of TK mRNA expression was observed in cells that expressed M1TK RNA while cells that expressed Δ M1TK RNA only exhibited a reduction of $9 \pm 3\%$, averaged over four experiments. Thus, it appears that cleavage of TK mRNA by M1TK RNA did indeed occur inside cells, with a subsequent reduction in the level of TK mRNA that could be translated. No products of the cleavage of TK mRNA were detected in our RNase protection assays presumably because these RNAs, which lack either a cap structure or a polyA sequence, are rapidly degraded by intracellular RNases.

Protein extracts of the infected cells were analyzed for the presence of the TK polypeptide. Polypeptides were transferred to two identical membranes and one was stained with a TK-specific antibody (anti-TK), described by Liu and Summers (1988), while the other was stained with a monoclonal antibody against the capsid protein ICP35 of HSV-1 (anti-ICP35), described by Liu and Roizman (1991). The expression of ICP35 serves as an internal control for the quantitation of expression of TK. The results of four independent experiments are summarized in Figure 13: a reduction of at least $76 \pm 5\%$, averaged over four experiments, in the level of TK protein was observed in cells that expressed M1TK RNA while a reduction of only $10 \pm 4\%$, averaged over four experiments, was seen in cells that expressed Δ M1TK RNA. The low level of inhibition found in cells that expressed Δ M1TK RNA was presumably due to an antisense effect.

These examples show that when M1 RNA is converted to an RIGS (M1GS RNA), it cleaves one particular substrate in a reaction governed by sequence specificity. A custom-designed RIGS specific for thymidine kinase sequence (M1TK RNA), cleaves the TK mRNA *in vitro* and can be stably expressed in mammalian cells and can reduce the level of expression of TK by $\geq 75\%$ when these cells are infected with HSV-1. The reduction in the level of expression of TK that can be ascribed to the antisense effect of the (internal) guide sequence is no more than 15% of the total inhibitory effect. Moreover, the high degree of sequence

specificity, which is governed by a guide sequence that hybridizes to a complementary sequence in the substrate, makes our construct suitable for use as a tool for targeting RNAs *in vivo*. The optimal length of an antisense sequence for *in vivo* targeting is about 13 nucleotides, as
5 discussed by Stein and Cheng, *Science* 261: 1004-1012 (1993). The extent of inactivation that we observed was very similar to that achieved when endogenous RNase P is used as the catalytic agent to cleave complexes of target TK mRNA and separate EGSs expressed from synthetic genes that have been stably incorporated into human cells in
10 tissue culture.

The activity of M1GS RNA was stimulated *in vitro* by C5 protein and mammalian proteins. Gold and Altman (1986) suggested that C5 protein and protein subunits of human RNase P might bind to homologous sequences and similar structures that are found in both M1
15 RNA and H1 RNA, the RNA component of RNase P from HeLa cells. The cleavage reactions by M1GS RNA in cells that contain proteins that bind RNase P catalytic RNA are expected to proceed at rates higher than those observed *in vitro*. In particular, higher rates are expected in the presence of additional, non-specific, RNA-binding proteins, RNA
20 chaperons, which are known to stimulate the activity of other RNA enzymes as indicated by Tsuchihashi *et al.*, *Science* 262: 99-102 (1993), Coetzee *et al.*, *Genes & Dev.* 8: 1575-1588 (1994), Bertrand and Rossi, *EMBO J.* 13: 2904-2912 (1994), and Herschlag *et al.*, *EMBO J.* 13: 2913-2924 (1994).

SEQUENCE LISTING

- (1) GENERAL INFORMATION:
- (i) APPLICANT: Yuan, Yan
Guerrier-Takada, Cecilia
Altman, Sidney
Liu, Fenyong
 - (ii) TITLE OF INVENTION: TARGETED CLEAVAGE OF RNA USING
RIBONUCLEASE P TARGETING AND CLEAVAGE SEQUENCES
 - (iii) NUMBER OF SEQUENCES: 65
 - (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Patrea L. Pabst
 - (B) STREET: 2800 One Atlantic Center
1201 West Peachtree Street
 - (C) CITY: Atlanta
 - (D) STATE: Georgia
 - (E) COUNTRY: USA
 - (F) ZIP: 30309-4530
 - (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
 - (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
 - (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: 08/207,547
 - (B) FILING DATE: 07-MAR-1994
 - (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Pabst, Patrea L.
 - (B) REGISTRATION NUMBER: 31,284
 - (C) REFERENCE/DOCKET NUMBER: YU112
 - (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (404)-873-8794
 - (B) TELEFAX: (404)-873-8795
- (2) INFORMATION FOR SEQ ID NO:1:
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 88 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: RNA (genomic)
 - (iii) HYPOTHETICAL: NO
 - (iv) ANTI-SENSE: NO
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
GAACAUUUUG AGGCAUUUCA GUCAGUUGGC CAAACUGAGC AGACUCUAAA 50
UCUGCNNNNN GAAGGUUCNN NNCCUUAUG CCUCACCA 88
- (2) INFORMATION FOR SEQ ID NO:2:
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 61 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
TAATACGACT CACTATAGAA CATTITGAGG CATTTCAGTC AGTTGGCCAA 50
ACTGAGCAGA C 61
(2) INFORMATION FOR SEQ ID NO:3:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 61 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
TGGTGAGGCA TGAAGGNNNN GAACCTTCNN NNNGCAGATT TAGAGTCTGC 50
TCAGTTTGGC C 61
(2) INFORMATION FOR SEQ ID NO:4:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 99 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: tRNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:
GAAUACACGG AAUUGGUGGG GUUCCCGAGC GGCCAAAGGG AGCAGACUCU 50
AAAUCUGCCG UCAUCGACUU CGAAGGUUCG AAUCCUUCCC CCACUGCCA 99
(2) INFORMATION FOR SEQ ID NO:5:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 71 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: tRNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:
GCCAAAGGGA GCAGACUCUA AAUCUGCCGU CAUCGACUUC GAAGGUUCGA 50
AUCCUUCCCC CACCACCAUC A 71
(2) INFORMATION FOR SEQ ID NO:6:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: tRNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:
GAAUACACGG AAUUGGUGGG GUUCCCGA 28
(2) INFORMATION FOR SEQ ID NO:7:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs

58

(B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: mRNA
 (iii) HYPOTHETICAL: NO
 (iv) ANTI-SENSE: NO
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:
 AUGGAACAUU UUGAGGCAUU UCAGUCAGUU UAA 33

(2) INFORMATION FOR SEQ ID NO:8:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 70 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: tRNA
 (iii) HYPOTHETICAL: NO
 (iv) ANTI-SENSE: NO
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:
 GGCCAAACUG AGCAGACUCU AAAUCUGCCG UCAUCGACUU CGAAGGUUCG 50
 AAUCCUUCAU GCCUCACCAU 70

(2) INFORMATION FOR SEQ ID NO:9:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 61 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: tRNA
 (iii) HYPOTHETICAL: NO
 (iv) ANTI-SENSE: NO
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:
 GGCCAAACUG AGCAGACUCU AAAUCGCAAA CGGAAGGUUC GUGCCCUUCA 50
 UGCCUCACCA U 61

(2) INFORMATION FOR SEQ ID NO:10:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 53 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: tRNA
 (iii) HYPOTHETICAL: NO
 (iv) ANTI-SENSE: NO
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:
 GGCCAAACUG ACGUCAUCGA CUUCGAAGGU UCGAAUCCUU CAUGCCUCAC 50
 CAU 53

(2) INFORMATION FOR SEQ ID NO:11:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 341 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: RNA
 (iii) HYPOTHETICAL: NO
 (iv) ANTI-SENSE: NO
 (vi) ORIGINAL SOURCE:
 (A) ORGANISM: Homo sapiens

- (x) PUBLICATION INFORMATION:
 (A) AUTHORS: Altman et al.
 (C) JOURNAL: Genomics
 (D) VOLUME: 18
 (F) PAGES: 418-422
 (G) DATE: 1993
 (K) RELEVANT RESIDUES IN SEQ ID NO:1: FROM 1 TO 341
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:
- (2) INFORMATION FOR SEQ ID NO:12:
- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 18 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:
GCCAAACTGA GCAGACTC 18
- (2) INFORMATION FOR SEQ ID NO:13:
- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 31 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:
GCGCGGTACC AAAAATGGTG AGGCATGAAG G 31
- (2) INFORMATION FOR SEQ ID NO:14:
- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 23 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:
GGCCGTAATA TCCAGCTGAA CGG 23
- (2) INFORMATION FOR SEQ ID NO:15:
- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 16 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:
TGGTGAGGCA TGAAGG 16
- (2) INFORMATION FOR SEQ ID NO:16:
- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 33 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:
TAATACGACT CACTATAGGC CAACTGAGCA GAC 33
(2) INFORMATION FOR SEQ ID NO:17:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 61 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: tRNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:
GGCCAAACUG AGCAGACUCU AAAUCGGCCC UUCGAAGGUU CGCCCCCUUC 50
AUGCCUCACC A 61
(2) INFORMATION FOR SEQ ID NO:18:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 62 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: tRNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:
GGCCAAACUG AGCAGACUCU AAAUCUGCAC GAGAGAAGGU UCGUGCCCUU 50
CAUGCCUCAC CA 62
(2) INFORMATION FOR SEQ ID NO:19:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 61 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: tRNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:
GGCCAAACUG AGCAGACUCU AACUGGCCU AACGAAGGUU CGCCCCCUUC 50
AUGCCUCACC A 61
(2) INFORMATION FOR SEQ ID NO:20:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 60 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: tRNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

61

GGCCAAACUG AGCAGACUCU AAUUGCCCA ACGAAGGUUC ACCCCCUUCA 50
 UGCCUCACCA 60

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 59 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: tRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GGCCAAACUG AGCAGACUCC AAUCCACCA AGAAGGUUCG UGCCCUUCAU 50
 GCCUCACCA 59

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 60 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: tRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GGCCAAACUG AGCAGACUCU AAUCCUCC CAGAAGGUUC GUGCCCUUCA 50
 UGCCUCACCA 60

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 61 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: tRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GGCCAAACUG AGCAGACUCU AAUCCGCCU ACGGAAGGUU CGCCCCCUUC 50
 AUGCCUCACC A 61

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 61 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: tRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GGCCAAACUG AGCAGACGCU AAUUCUACCC CGUGAAGGUU CGUCCCCUUC 50
 AUGCCUCACC A 61

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 61 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: tRNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:
GGCCAAACUG AGCAGACUCU AAUUUGCCA CCAGAAGGUU CGCCCCCUUC 50
AUGCCUCACC A 61
(2) INFORMATION FOR SEQ ID NO:26:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 61 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: tRNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:
GGCCAAACUG AGCAGACUCA AAUCUGGCCA UUCGAAGGUU CGCCCCCUUC 50
AUGCCUCACC A 61
(2) INFORMATION FOR SEQ ID NO:27:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 60 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: tRNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:
GGCCAAACUG AGCAGACUCU AAAUCGCGAGU GUGAAGGUUC GUGCCCUUCA 50
UGCCUCACCA 60
(2) INFORMATION FOR SEQ ID NO:28:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 61 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: tRNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:
GGCCAAACUG AGCAGACUCU AAAUCAGCGC GUGGAAGGUU CGUGCCCUUC 50
AUGCCUCACC A 61
(2) INFORMATION FOR SEQ ID NO:29:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 61 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: tRNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:
GGCCAAACUG AGCAGACUCU AAAUCGGCCG CACGAAGGUU CGCCCCCUUC 50
AUGCCUCACC A 61

- (2) INFORMATION FOR SEQ ID NO:30:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 61 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: tRNA
 (iii) HYPOTHETICAL: NO
 (iv) ANTI-SENSE: NO
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:
GGCCAAACUG AGCAGACACU AAUUUGCAC GAGGAAGGUU CGCCCCCUUC 50
AUGCCUCACC A 61
- (2) INFORMATION FOR SEQ ID NO:31:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 61 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: tRNA
 (iii) HYPOTHETICAL: NO
 (iv) ANTI-SENSE: NO
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:
GGCCAAACUG AGCAGACCCU AAUUCUGCCC CCGGAAGGUU CGUGCCCUUC 50
AUGCCUCACC A 61
- (2) INFORMATION FOR SEQ ID NO:32:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 24 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (iii) HYPOTHETICAL: NO
 (iv) ANTI-SENSE: NO
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:
GCCAAACTGA CGTCATCGAC TTCG 24
- (2) INFORMATION FOR SEQ ID NO:33:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 16 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (iii) HYPOTHETICAL: NO
 (iv) ANTI-SENSE: NO
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:
AACAGCTATG ACCATG 16
- (2) INFORMATION FOR SEQ ID NO:34:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 59 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (iii) HYPOTHETICAL: NO
 (iv) ANTI-SENSE: NO

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:
 GTAATACGAC TCACTATAGG CCAAACGTGAG CAGACTCTAA ATCTGCAAAC 50
 GGAAGGTTC 59
- (2) INFORMATION FOR SEQ ID NO:35:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 47 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: tRNA
 (iii) HYPOTHETICAL: NO
 (iv) ANTI-SENSE: NO
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:
 ATGGCTTCGT ACCCCTGCCA TCAACACGCG TCTGCGTTTCG ACCAGGC 47
- (2) INFORMATION FOR SEQ ID NO:36:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 71 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: tRNA
 (iii) HYPOTHETICAL: NO
 (iv) ANTI-SENSE: NO
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:
 GGTAAACGTC GGACAGACTC TAAATCTGTT GCGGTCTCCG CGCGCAGGTT 50
 CAAATCCTGC CGCAGACGTT T 71
- (2) INFORMATION FOR SEQ ID NO:37:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 71 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: tRNA
 (iii) HYPOTHETICAL: NO
 (iv) ANTI-SENSE: NO
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:
 GGTAAACGTC GGACAGACTC TAAATCTGTT GCGGTCTCCG CGCGCAGGTT 50
 GAAATCCTGC CGCAGACGTT T 71
- (2) INFORMATION FOR SEQ ID NO:38:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 54 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: tRNA
 (iii) HYPOTHETICAL: NO
 (iv) ANTI-SENSE: NO
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:
 GGTAAACGTC GGTGCGGTCT CCGCGCGCAG GTTCAAATCC TGCCGCAGAC 50
 GTTT 54
- (2) INFORMATION FOR SEQ ID NO:39:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 62 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: tRNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:
GGCCAAACUG AGCAGACUCU AAAUCUGCAA ACGGAAGGUU CGUGCCCUUC 50
AUGCCUCACC AU 62

(2) INFORMATION FOR SEQ ID NO:40:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 377 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: RNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:
GAAGCUGACC AGACAGUCGC CGCUUCGUCG UCGUCCUCUU CGGGGGAGAC 50
GGGCGGAGGG GAGGAAAGUC CGGGCUCCAU AGGGCAGGGU GCCAGGUAAC 100
GCCUGGGGGG GAAACCCACG ACCAGUGCAA CAGAGAGCAA ACCGCCGAUG 150
GCCCCGCGCAA GCGGGAUCAG GUAAGGGUGA AAGGGUGCGG UAAGAGCGCA 200
CCGCGCGGCU GGUAACAGUC CGUGGCACGG UAAACUCCAC CCGGAGCAAG 250
GCCAAAUAGG GGUUCAUAAG GUACGGCCCG UACUGAACCC GGUUAGGCUG 300
CUUGAGCCAG UGAGCGAUUG CUGGCCUAGA UGAAUGACUG UCCACGACAG 350
AACCCGGCUU AUCGGUCAGU UUCACCU 377

(2) INFORMATION FOR SEQ ID NO:41:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:
TAATACGACT CACTATAG 18

(2) INFORMATION FOR SEQ ID NO:42:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:
CGCAGACGGT CCTATAGTGA GTCGTATTA 29

(2) INFORMATION FOR SEQ ID NO:43:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 63 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA
(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:
 ACCGCGCAGC CTGGTCGAAC GCAGACGCGT GTTGATGGCA GGGGTCTATA 50
 GTGAGTCGTA TTA 63

(2) INFORMATION FOR SEQ ID NO:44:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 29 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA
 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:
 ATGCCTCGGT CCTATAGTGA GTCGTATTA 29

(2) INFORMATION FOR SEQ ID NO:45:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 32 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA
 (iii) HYPOTHETICAL: NO
 (iv) ANTI-SENSE: NO
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:
 GTGGTGTCTG CGTTCGACCA GCTATGACCA TG... 32

(2) INFORMATION FOR SEQ ID NO:46:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 31 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA
 (iii) HYPOTHETICAL: NO
 (iv) ANTI-SENSE: NO
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:
 GTGGTGTCTG CGTTCGACCA GTATGACCAT G 31

(2) INFORMATION FOR SEQ ID NO:47:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 28 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA
 (iii) HYPOTHETICAL: NO
 (iv) ANTI-SENSE: NO
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:
 GTGGTGTCTG CGTTCGACTA TGACCATG 28

(2) INFORMATION FOR SEQ ID NO:48:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:
GTGGTGTCTG CGTTCTATGA CCATG 25
- (2) INFORMATION FOR SEQ ID NO:49:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: DNA
 - (iii) HYPOTHETICAL: NO
 - (iv) ANTI-SENSE: NO
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:
GTGGTGTCTG TATGACCATG 20
- (2) INFORMATION FOR SEQ ID NO:50:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 28 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: DNA
 - (iii) HYPOTHETICAL: NO
 - (iv) ANTI-SENSE: NO
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:
GTGGTGAGGC ATTCAGTTA TGACCATG 28
- (2) INFORMATION FOR SEQ ID NO:51:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 12 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: RNA
 - (iii) HYPOTHETICAL: NO
 - (iv) ANTI-SENSE: NO
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:
AUGCCUCACC AC 12
- (2) INFORMATION FOR SEQ ID NO:52:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 12 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: RNA
 - (iii) HYPOTHETICAL: NO
 - (iv) ANTI-SENSE: NO
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:
CGCAGACACC AC 12
- (2) INFORMATION FOR SEQ ID NO:53:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 12 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: RNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:
GGACCGAGGC AU 12
(2) INFORMATION FOR SEQ ID NO:54:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: RNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:
GGACCGUCUG CG 12
(2) INFORMATION FOR SEQ ID NO:55:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: RNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:
GUCUGCGUUC GA 12
(2) INFORMATION FOR SEQ ID NO:56:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: RNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:
GAGGCAUUUC AGU 13
(2) INFORMATION FOR SEQ ID NO:57:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 46 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:
GACCCCTGCC ATCAACACGC GTCTGCGTTC GACCAGGCTG CGCGGT 46
(2) INFORMATION FOR SEQ ID NO:58:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: RNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:
AGCCUGGUCG AACGCAGACA CAAAAA 26
(2) INFORMATION FOR SEQ ID NO:59:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: RNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:
AGCCUGGUCG AACGCAGACA CCAC 24
(2) INFORMATION FOR SEQ ID NO:60:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: RNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:
CUGGUCGAAC GCAGACACCA C 21
(2) INFORMATION FOR SEQ ID NO:61:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: RNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:
GUCGAACGCA GACACCAC 18
(2) INFORMATION FOR SEQ ID NO:62:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: RNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:
GAACGCAGAC ACCAC 15
(2) INFORMATION FOR SEQ ID NO:63:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

70

(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: RNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:
CAGACACCAC 10

(2) INFORMATION FOR SEQ ID NO:64:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: RNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:
GUCGAACGCA GACACCAC 18

(2) INFORMATION FOR SEQ ID NO:65:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: RNA
(iii) HYPOTHETICAL: NO
(iv) ANTI-SENSE: NO
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:
CUGGUCGAAC GCAGACACCA C 21

We claim:

1. An isolated oligonucleotide molecule comprising an external guide sequence wherein the external guide sequence comprises
a recognition sequence complementary to a targeted sequence in a target RNA molecule, and
an RNase P binding sequence,
wherein the RNase P binding sequence comprises a nucleotide sequence base pairing with itself to form a structure similar to the T stem and loop of a precursor tRNA, wherein the RNase P binding sequence does not form structure similar to all or portions of the variable stem, the variable loop, the anticodon stem and the anticodon loop of a precursor tRNA, and
wherein the external guide sequence promotes RNase P catalytic RNA-mediated cleavage of the target RNA molecule.
2. The oligonucleotide molecule of claim 1 further comprising an RNase P catalytic sequence,
wherein the oligonucleotide molecule cleaves the target RNA.
3. The oligonucleotide molecule of claim 2 wherein the RNase P catalytic sequence is an H1 sequence or an M1 sequence.
4. The oligonucleotide molecule of claim 2 wherein the recognition sequence comprises an A recognition arm and a D recognition arm, wherein the A recognition arm is located at the 3' end of the RNase P binding sequence and the D recognition arm is located at the 5' end of the RNase P binding sequence.
5. The oligonucleotide molecule of claim 4 wherein the RNase P catalytic sequence is linked to the 3' end of the A recognition arm.
6. The oligonucleotide molecule of claim 4 wherein the RNase P catalytic sequence is linked to the 5' end of the D recognition arm.
7. The oligonucleotide molecule of claim 4 wherein the RNase P binding sequence consists of
a nucleotide sequence base pairing with itself to form a structure similar to the T stem and loop of a precursor tRNA.

8. The oligonucleotide molecule of claim 2 wherein the RNase P catalytic sequence comprises the sequence of a naturally occurring RNase P catalytic RNA.

9. The oligonucleotide molecule of claim 2 wherein the oligonucleotide molecule is selected by

- randomizing a section of the sequence of the starting oligonucleotide molecule;
- selecting for a subpopulation of the randomized sequences for their ability to efficiently cleave the target RNA;
- amplifying those sequences cleaving more efficiently than the starting oligonucleotide molecule; and
- repeating the selection and amplification steps.

10. A composition for promoting cleavage of a target RNA wherein the composition comprises the oligonucleotide molecule of claim 2 in a pharmaceutically acceptable delivery system.

11. The composition of claim 10 wherein the pharmaceutically acceptable delivery system is selected from the group consisting of liposomes, virosomes, microspheres and microcapsules.

12. The composition of claim 10 wherein the pharmaceutically acceptable delivery system is selected from the group consisting of carriers suitable for topical, subcutaneous, parenteral, and enteral administration.

13. A composition for promoting cleavage of a target RNA wherein the composition comprises an engineered expression vector encoding the oligonucleotide molecule of claim 2.

14. The composition of claim 13 wherein the engineered expression vector is a viral vector selected from the group consisting of retroviral vectors, adeno-associated viral vectors and Epstein-Barr viral vectors.

15. A method for cleaving a target RNA comprising

- bringing into contact, under conditions that promote RNase P cleavage, the target RNA and an oligonucleotide molecule comprising an external guide sequence wherein the external guide sequence comprises

a recognition sequence complementary to a targeted sequence in a target RNA molecule, and

an RNase P binding sequence,

wherein the RNase P binding sequence comprises a nucleotide sequence base pairing with itself to form a structure similar to the T stem and loop of a precursor tRNA, wherein the RNase P binding sequence does not form structure similar to all or portions of the variable stem, the variable loop, the anticodon stem and the anticodon loop of a precursor tRNA, and

wherein the external guide sequence promotes RNase P catalytic RNA-mediated cleavage of the target RNA molecule.

16. The method claim 15 wherein the oligonucleotide molecule further comprises

an RNase P catalytic sequence,

wherein the oligonucleotide molecule cleaves the target RNA.

17. The method of claim 16 wherein the step of bringing into contact is accomplished by administering to a patient or cells from a patient the oligonucleotide molecule, and

wherein the oligonucleotide molecule is in a pharmaceutically acceptable delivery system.

18. A method for selecting a population of RNase P internal guide sequences that cleave a target RNA with increased efficiency over a starting RNase P internal guide sequence comprising

randomizing a section of the starting RNase P internal guide sequence;

selecting for a subpopulation of the randomized sequences for their ability to efficiently cleave the target RNA;

amplifying those sequences cleaving more efficiently than the starting RNase P internal guide sequence; and

repeating the selection and amplification steps.

19. An isolated oligonucleotide molecule comprising an external guide sequence wherein the external guide sequence comprises
a recognition sequence complementary to a targeted sequence in a target RNA molecule, and
an RNase P binding sequence;
wherein the external guide sequence promotes RNase P catalytic RNA-mediated cleavage of the target RNA molecule, and
wherein at least one nucleotide in the oligonucleotide molecule is selected from the group consisting of chemically modified nucleotides and chemically unmodified deoxyribonucleotides.

20. The oligonucleotide molecule of claim 19 wherein the recognition sequence comprises an A recognition arm and a D recognition arm, wherein the A recognition arm is located at the 3' end of the external guide sequence and the D recognition arm is located at the 5' end of the external guide sequence.

21. The oligonucleotide molecule of claim 20 wherein
the A recognition arm comprises a nucleotide sequence including at least seven nucleotides complementary to and base pairing with the substrate immediately 3' to a site in the substrate to be cleaved to form a structure similar to an aminoacyl acceptor stem,

the RNase P binding sequence comprises a nucleotide sequence base pairing with itself to form a structure similar to the T stem and loop of a precursor tRNA, and

the D recognition arm comprises a nucleotide sequence including at least three nucleotides complementary to and base pairing with the substrate.

22. The oligonucleotide molecule of claim 21 wherein the RNase P binding sequence comprises

a nucleotide sequence base pairing with itself to form a structure similar to the T stem and loop of a precursor tRNA, and

a nucleotide sequence forming structure similar to all or portions of the variable stem, the variable loop, the anticodon stem and the anticodon loop of a precursor tRNA.

23. A composition for promoting cleavage of a target RNA molecule wherein the composition comprises the oligonucleotide molecule of claim 19 in a pharmaceutically acceptable delivery system.

24. The composition of claim 23 wherein the pharmaceutically acceptable delivery system is selected from the group consisting of liposomes, virosomes, microspheres and microcapsules.

25. The oligonucleotide molecule of claim 19 wherein one or more of the 2' hydroxyl groups of the ribonucleotides are replaced with a chemical group selected from the group consisting of hydrogen, an O-alkyl group, an amino group, and fluorine,

wherein one or more of the phosphate linking groups are replaced with a linking group selected from the group consisting of methyl phosphonate and phosphorothioate, and

wherein replacement of one or more of the 2' hydroxyl groups increases resistance of the external guide sequence to nucleases.

26. The oligonucleotide molecule of claim 19 wherein one or more of the 2' hydroxyl groups of the ribonucleotides are replaced with hydrogen or a methoxy group; and

wherein one or more of the phosphate linking groups are replaced with phosphorothioate.

27. The oligonucleotide molecule of claim 19 further comprising an RNase P binding sequence,

wherein the oligonucleotide molecule cleaves the target RNA.

28. An DNA molecule that encodes an RNA molecule comprising an external guide sequence wherein the external guide sequence comprises a recognition sequence complementary to a targeted sequence in a target RNA molecule, and

an RNase P binding sequence,

wherein the RNase P binding sequence comprises a nucleotide sequence base pairing with itself to form a structure similar to the T stem and loop of a precursor tRNA, wherein the RNase P binding sequence does not

form structure similar to all or portions of the variable stem, the variable loop, the anticodon stem and the anticodon loop of a precursor tRNA, and wherein the external guide sequence promotes RNase P catalytic RNA-mediated cleavage of the target RNA molecule.

29. The DNA molecule of claim 28 wherein the RNA molecule further comprises

an RNase P catalytic sequence,
wherein the RNA molecule cleaves the target RNA.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 95/02816

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/52 C12N9/00 A61K31/70 C12Q1/68 C07H21/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N A61K C12Q C07H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JOURNAL OF MOLECULAR BIOLOGY, vol. 226, 1992 pages 399-409, PERREAULT, J.-P. & ALTMAN, S. 'Important 2'-hydroxyl groups in model substrates for M1 RNA, the catalytic RNA subunit of RNase P from Escherichia coli' see page 400, right column, paragraph 3 see figure 1A	1, 15, 19-21, 25, 27, 28
Y	see page 407 - page 408 --- -/--	1-29

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

1 August 1995

Date of mailing of the international search report

16. 08. 95

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,
Fax (+ 31-70) 340-3016

Authorized officer

Andres, S

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 95/02816

C(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	<p>NUCLEIC ACIDS RESEARCH, vol. 21, no. 20, 11 October 1993 OXFORD GB, pages 4685-4689, KIKUCHI, Y. ET AL. 'Artificial self-cleaving molecules consisting of a tRNA precursor and the catalytic RNA of RNase P' see the whole document ---</p>	<p>1-8, 15-17, 19-21, 27,29</p>
Y	<p>SCIENCE (WASHINGTON D C) 263 (5151). 1269-1273, 4 March 1994 YUAN, Y. ET AL. 'Selection of guide sequences that direct efficient cleavage of mRNA by human ribonuclease P.' cited in the application see the whole document ---</p>	<p>9,18</p>
Y	<p>WO-A-93 22434 (YALE UNIVERSITY, USA) 11 November 1993 cited in the application see page 6, line 13 - page 7, line 6 see page 10, line 4 - line 20 see page 13, line 26 - page 20, line 10 see examples see figure 5B 6B 6C ---</p>	<p>10-14, 22-29</p>
A	<p>SCIENCE (WASHINGTON D C) 249 (4970). 1990. 783-786, FORSTER, A. ET AL. 'EXTERNAL GUIDE SEQUENCES FOR AN RNA ENZYME.' see figure 1B ---</p>	<p>1-29</p>
A	<p>PROC NATL ACAD SCI U S A 89 (17). 1992. 8006-8010, YUAN, Y. ET AL. 'TARGETED CLEAVAGE OF MRNA BY HUMAN RNASE P.' cited in the application see page 8008, left column, line 17 - line 19 ---</p>	<p>1-29</p>
A	<p>NUCLEIC ACIDS RESEARCH, vol. 21, no. 1, 1993 OXFORD GB, pages 21-26, GAUR, R. & KRUPP, G. 'Modification interference approach to detect ribose moieties important for the optimal activity of a ribozyme' see the whole document ---</p>	<p>19-26</p>

1

-/--

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 95/02816

C(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	<p>EMBO (EUROPEAN MOLECULAR BIOLOGY ORGANIZATION) JOURNAL 14 (1). 159-168, 3 January 1995 YUAN, Y. & ALTMAN, S. 'Substrate recognition by human RNase P: Identification of small, model substrates for the enzyme.' see the whole document -----</p>	<p>1,9,15, 18-21,28</p>

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 95/02816

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A-9322434	11-11-93	AU-B- 4119893	29-11-93
		EP-A- 0638121	15-02-95

INTERNATIONAL SEARCH REPORT

International Application No.

PCT/US95/02816

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.: 15-17
because they relate to subject matter not required to be searched by this Authority, namely:
Remark: Although claims 15-17 (as far as in vivo methods are concerned) are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.